Universität
Bremen

Learning to Teach Through Proving:

In-service Primary School Teachers' Understanding and Use of Proving while engaged in Proof-Based Teaching

by

Estela A. Vallejo-Vargas

University of Bremen

Bremen, Germany

August 2023

Learning to Teach Through Proving:

In-service Primary School Teachers' Understanding and Use of Proving while engaged in Proof-Based Teaching

A dissertation submitted to Faculty 3, Mathematics and Computer Science, of the University of Bremen in fulfillment of the requirements for the degree of Dr. rer. nat.

Submitted by: Estela A. Vallejo-Vargas

Submission Date: August 2023

Supervisor: Prof. Dr. David A. Reid

Graduation Colloquium Date:

November 22nd, 2023.

Supervisor:

Prof. Dr. David A. Reid

Evaluators:

Prof. Dr. David A. Reid

Prof. Dr. Angelika Bikner-Ahsbahs

Acknowledgements

I want to begin expressing my special and deep gratitude to my supervisor and friend Prof. Dr. David Reid. I remember the first time I met Prof. Reid at the 38th Conference of the International Group for the Psychology of Mathematics Education in 2014. Prof. Reid was kind and took the time to listen to my ideas. He is a high-achieving researcher and well-recognized scholar in our field and I highly value his humbleness and the fact that he is always willing to listen to everyone. It has been a privilege and invaluable gift to work with him and to learn from him. I will never forget our meetings, especially those in Starbucks. He used to say that I had a "skeptical face" and I found his comment very accurate and a compliment. Prof. Reid supported my "crazy" ideas and encouraged me to pursue them. Because he was always close despite the distance, I want to thank Prof. Reid. This has been a long journey and I am profoundly grateful because achieving this goal would not have been possible without his guidance and support. To him and his family (Prof. Knipping, Sophia and Lili), for welcoming me to their home, thank you very much.

My gratitude also goes to Prof. Dr. Angelika Bikner-Ahsbahs, who always had a positive attitude about things and was close enough so that I felt comfortable talking to her. I appreciate the time Prof. Bikner-Ahsbahs took to provide her feedback. It has been an honor to meet such a wonderful woman who is always full of good ideas.

I also want to thank Prof. Uldarico Malaspina, my master's supervisor. Prof. Malaspina pushed me to pursue a PhD after I finished my master studies. We shared wonderful moments not only in Peru, but also as we had the chance to participate in different academic events abroad. I am deeply grateful to him for all the stories he told me and his friendship.

Thanks also to the University of Bremen and the MAL-project that supported my stay in Bremen. Without their financial support my efforts to work with interesting and knowledgeable researchers would not have been possible.

Thanks to all the good friends I met in Bremen. In Bremen I felt at home because of them. To my dear friend Rubina, who is an example of perseverance and dedication as she herself struggled in her attempts to accomplish her goals in Germany. My "latina" looking friend and I shared so many endless interesting conversations, which I will always remember. Thanks to her I became aware of the reality about her hometown (Xinjiang) and about the world in general. Rubina, my twin, always made me laugh, even in very difficult moments. Thanks to my dear friend Noah. It was a beautiful coincidence to have met someone who had spent one year in Trujillo, which is a northern city in Peru where I grew up. He was there for me when I felt that I wanted to quit, but also to enjoy memorable moments in Bremen. My Iranian friend Habibeh and her family always treated me as if I was one of them. Nele and her husband Gideon enjoyed many meals with me at my place and helped me when I needed. Thanks for all the laughter and all the love I received from all of them. I value our friendship and I feel lucky to have you in my life as my friends.

I also want to express my gratitude to my family who supported all of my academic endeavors. Once I returned to my country (Peru) they helped me try to keep focused despite the hard times that the pandemic forced us to face. It has been a long journey and I value their trust, patience and love.

Thanks to the teachers who participated in my two studies: Manuela, Priscila, Irene, Amilta, Andrea, Gessenia and Lizbeth, and their wonderful students. I am in debt with Ursula Asmad who gave me all the support I could have needed to develop my research in the schools. The time I spent in the schools with the teachers and their students was a wonderful learning experience and joy.

Last but not least, I want to thank God because I had the big opportunity to accomplish one of my dreams. With this experience I have learned that this world still has people who selflessly give away their time to help others. I hope I can give back all of what these wonderful people have done for me and contribute to the mathematics education community around the globe and in particular the mathematics education of my country, Peru, which I am proud of.

Abstract

This study investigates the effects of an intervention that has a focus on the development of in-service teachers' proof-related assumptions and the way their understanding of the nature of proving is observable during their teaching. Three in-service teachers of a private school from Peru participated in the study. They were first engaged in an exploratory interview that consisted of two parts and gave access to the teachers' initial proof-related assumptions. Based on the observations during the exploratory interview, refinements were made to an initial design of an intervention that aimed at supporting the teachers' change from non-mathematical proof-related assumptions to the formation of mathematical ones. The development of the intervention was part of a two-year design-based research. Here the focus is on the second research cycle which is a refinement of the first cycle. The teachers used the insights they gained during the intervention to teach in their classrooms. Their teaching made visible their use of mathematical assumptions they had changed during the intervention and some new unexpected mathematical assumptions.

This study does not only show the way these three teachers developed their own understanding. It shows that although the teachers might use mathematically-aligned assumptions, their warrants might not be mathematical. More importantly it shows the close relation among the teachers' assumptions and the way it influences their development. It underscores the need for research that emphasizes a global view of individuals' assumptions so that a real understanding of their development can be achieved.

This study identifies the features of the intervention that may explain changes in the three teachers' proof-related assumptions. Those features were extrapolated to make recommendations of what an intervention with similar goals should include. Suggestions for future research and some reflections are also included.

Table of Contents

Chapter 1: Introduction and Research Questions

"If problem solving is the 'heart of mathematics,' then proof is its soul."

(Schoenfeld, 2011, p. xii)

The goals of my research are to (1) design an intervention with a focus on supporting the development of primary school teachers' proof-related understandings and, notably, their understanding of the nature of proving, (2) investigate the effects the intervention had on the teachers' abilities to engage in proof-related activities and how it is reflected while they teach in their classrooms in a proof-based way, and (3) identify the features of the intervention that may explain changes in the teachers' proof-related understandings.

The remarkable role that proof plays in mathematics education has been broadly recognized (see e.g., Ball & Bass, 2003; Harel & Sowder, 2007; National Council of Teachers of Mathematics [NCTM], 2000; G. J. Stylianides, 2008; G. J. Stylianides & A. J. Stylianides, 2008). In particular, G. J. Stylianides and A. J. Stylianides (2008) stated that there were three main reasons to explain the increasing emphasis on proof in mathematics education.

> First, proof is fundamental to doing mathematics—it is the basis of mathematical understanding and is essential for developing, establishing, and communicating mathematical knowledge … Second, students' proficiency in proof can improve their mathematical proficiency more broadly because proof is "involved in all situations where conclusions are to be reached and decisions to be made" … Third, several researchers … have identified students' abrupt introduction to proof in high school as a possible explanation for the many difficulties that secondary school … and university students … face with proof, thereby proposing that students engage with proof in a coherent and systematic way throughout their schooling. (p. 104)

Related to the first reason, Schoenfeld (2011) pointed out the important role that proving plays in the mathematical activity, which is mainly related to the work of mathematicians.

> This dualism of exploration and confirmation, the lifeblood of mathematics, is the everyday work of every mathematician. Yet, the vitality of proving as an essential part of doing mathematics—of exploring, conjecturing, and subjecting one's conjectures to rigorous testing—is little understood outside the mathematical community. (p. xii)

Given its relevance, there is a worldwide tendency to increase the role of mathematical proofs in school programs (e.g., Ball & Bass, 2003; Ball, Hoyles, Jahnke & Movshovitz-Hadar, 2002; Hanna & Jahnke, 1996; Mariotti, 2006; NCTM, 2000; A. J. Stylianides, 2007; G. J. Stylianides, 2016; Yackel & Hanna, 2003). In the last decades, the amount of research on the teaching and learning of mathematical proof has substantially increased (for overviews see Blanton, et al., 2011; Hanna, 2000; Hanna & de Villiers, 2012; Mariotti, 2006; Reid & Knipping, 2010; G. J. Stylianides, A. J. Stylianides & Weber, 2017). Argumentation, including mathematical proofs, is a competency assessed by the PISA tests (OECD, 2009, p. 106), and "Proof and Reasoning" is a focus of the influential *Principles and Standards for School Mathematics* (NCTM, 2000). The NCTM *Standards* document emphasized that students' mathematical education throughout the school levels should prepare them progressively "[t]o recognize reasoning and proof as fundamental aspects of mathematics, make and investigate mathematical conjectures, develop and evaluate mathematical arguments and proofs, and select and use various types of reasoning and methods of proof" (p. 56).

The value of mathematical proof in the classroom reflects its role in mathematics. As Recio (2002) stated:

> The educational value of mathematical proof appears as part of a more general utility, which is to learn to reason mathematically. To reason operationally, to solve problems and to justify the generalized compliance of the mathematical propositions used in these problem-solving processes, helps students to build an intelligent, logic, mathematical edifice and not just a functional one. (p. 38, my translation)

This process of building mathematics is central to constructivist mathematics education.

> Constructivist mathematical education tries to show that, with something we already know, with some help and with a good management of the available information, we are capable of solving problems. Along the way, we learn through creating mathematics what our mathematics will be. It might not be groundbreaking, worldwide knowledge, but we will have carried out a process that is almost identical to that of a professional mathematician. (Albertí, 2011, p. 31, my translation)

Despite all the important issues previously pointed out, in most teaching approaches mathematical proof is seen as a subject of study, which does not reflect its role in mathematics as a way to construct mathematical knowledge. Reid (2011) suggested the need for a new approach, "Proof Based Teaching" (PfBT), in which "we see proof as fundamental to mathematics as a way to develop understanding of mathematical concepts, and as a way to discover new and significant mathematical knowledge" (p. 28). Reid made an analogy with the role of problem solving in mathematics education. Problem solving was also seen for a long time as a subject of study, but this perspective did not make anybody more capable of solving problems. More recently, problem solving has come to be seen as a means through which it is possible to teach mathematics.

> Today proof teaching and research is in a similar state to problem solving in the 1980s. Much research is being done, but proof is still mostly seen as a topic to be taught. In the case of problem solving we have moved towards seeing it as a way of teaching, and there are early hints that something like proof based teaching, akin to problem based teaching, might emerge in the next decades. It will be very interesting to see. (Reid, 2011, p. 28)

In addition, developing proof-related skills entails several challenges. Research has shown that students at different educational levels have difficulty engaging in proving (e.g., Harel & Sowder, 2007; Reid & Knipping, 2010; G. J. Stylianides, A. J. Stylianides & Weber, 2017). Concerning the reason why some students have difficulties when they get involved in reasoning and proving, G. J. Stylianides, A. J. Stylianides and Shilling-Traina (2013) pointed out that "students' difficulties with reasoning-and-proving should be attributed less to cognitive-developmental constraints and more to the fact that prevalent classroom instruction places limited emphasis on reasoning-and-proving" (p. 1464). At least partially that is the result of pre-service and in-service school teachers facing similar difficulties, both in engaging in proving and understanding the nature of proving and its role in mathematics and mathematics learning (e.g., Barkai, Tsamir, Tirosh & Dreyfus, 2002; Knuth, 2002a, 2002b; Martin & Harel, 1989). However, changing those forms of reasoning and thinking is not easy.

> Trying to change thinking habits, especially ones that have been ingrained over a period of years, is a very difficult task… in order to learn a complex process such as proof and disproof, effective integration of new modes of thought with pre-existing contradictory modes is a major undertaking. It is not surprising that easy solutions have not yet been discovered. (Epp, 2003, p. 893)

A notable challenge within this path is the lack of preparation teachers have to foster their students' proof-related skills. As A. J. Stylianides and Ball (2008, p. 309) noted, "the growing appreciation of the importance of making proving central to all students' mathematical experiences and as early as the elementary grades … places significant demands on teachers' knowledge about proof". Hence, developing proof-related abilities in students involves that the teachers understand the processes entailed in reasoning and proving. In A. J. Stylianides and Ball's (2008) terms, "[u]nless teachers have good understanding of proof, we cannot expect that they will be able to effectively promote proving among their students" (p. 309).

In the same line, G. J. Stylianides et al. (2013) noted that,

> reasoning-and-proving tends to have a marginal place in elementary school classrooms. This situation can be partly attributed to the fact that many (prospective) elementary teachers have (1) weak mathematical (subject matter) knowledge about reasoning-and-proving and (2) counterproductive beliefs about its teaching. (p. 1463)

Among the counterproductive beliefs these teachers have about proof it can be highlighted the beliefs that: (a) the main role of proof is verification, (b) the favorable mathematics topic for students to work with proof is Euclidean Geometry, and (c) it is difficult to promote proving among young students (e.g., Healy & Hoyles, 2000; Knuth, 2002a, 2002b; Cabassut, et al., 2012).

In fact, the bulk of research related to reasoning and proving in mathematics education has mainly focused on identifying the difficulties that individuals face when engaged in dis/proving or when evaluating others' arguments. Similarly, research has paid great attention to exploring individuals' (in particular, teachers') conceptions about proof (e.g., Healy & Hoyles, 2000; Knuth, 2002a, 2002b; for a review, see Ko, 2010). Even though such studies are relevant in order to guide the design of interventions that aim at improving individuals' poor proving skills, there is still much to do in terms of finding ways to remediate those difficulties and the poor conceptions and beliefs that individuals share in such a context (G. J. Stylianides et al., 2017). G. J. Stylianides and A. J. Stylianides (2017) pointed out that "less emphasis has been placed on acting upon such problems to generate possible solutions through the design, implementation, and evaluation of research-based interventions in mathematics classrooms" (p. 120). Even though there are some examples of research-based interventions with proof as a focus (e.g., Buchbinder & McCrone, 2020, 2022; G. J. Stylianides & A. J. Stylianides, 2009), their number is still small. A special 2017 issue in Educational Studies in Mathematics has been published to try to address this lack. In it, G. J. Stylianides and A. J. Stylianides (2017) called for developing research that aims at the design of classroom-based interventions at the elementary school level with a focus on key and persistent problems of students' learning in the area of proof; for example, the assumptions that empirical arguments suffice to prove a generalization, a single counterexample is insufficient to refute a false mathematical generalization, or a conditional statement and its converse are equivalent.

Some researchers have suggested directions on the aspects of proof to which interventions should pay attention (e.g., Buchbinder & McCrone, 2020, 2022; Durand-Guerrier, et al., 2012; Epp, 2003, 2009b; A. J. Stylianides, 2011; A. J. Stylianides & Ball, 2008; G. J. Stylianides & A. J. Stylianides, 2009, 2014, 2017; G. J. Stylianides, et al., 2013). For instance, Epp (2003) recommended including some basic principles of logical reasoning to support the development of proof-related understandings.

> Another important reason for discussing logic explicitly is to help prepare the next generation of teachers. Unfortunately, at least in the United States, a large number of K-12 teachers have only a weak command of the principles of logical reasoning. It is simply not possible for such teachers to promote effectively their students' reasoning development … when they themselves do not have a feeling for what is or is not a valid deduction or what it means for statements of various forms to be true or false. (p. 894)

Besides considering knowledge about the logico-linguistic structure of proof, A. J. Stylianides and Ball (2008) claimed that such knowledge should be complemented with knowledge of situations for proving.

Some researchers have noted that these suggestions should involve the teachers' understanding of the nature and role of proof-related concepts (e.g., Chazan, 1990; Jones, 1997; Knuth, 2002a, 2002b). Notably, the teachers should engage themselves in proof-related discussions before they attempt to engage others in similar discussions. Knuth (2002b) recommended that "[a] starting point toward helping teachers adopt and implement such a perspective may be to engage teachers in explicit discussions about proof" (p. 83).

In terms of my prior experience, my masters research (Vallejo-Vargas, 2012) was focused on designing, teaching and reflecting on an instructional sequence to promote 12-year old students' construction of knowledge about divisibility. In the Peruvian National Curriculum this was the age at which students first learned this mathematical content. However, in my work with these students I noticed they had already seen divisibility before in their schooling, although in a traditional way. For this reason, my research aims changed to leading the students to construct new knowledge by making and proving conjectures, and to justify knowledge they have acquired from their teacher in order to make sense of divisibility.

The work of Ordoñez-Montañez (2014) is a continuation of my masters' research and provides a clear example of what Reid (2011) called Proof-Based Teaching (Vallejo-Vargas & Ordoñez-Montañez, 2015). In this work, for which I was the supervisor, Ordoñez-Montañez showed how third graders were capable of constructing their own knowledge of division and divisibility of natural numbers from three key notions: fair, whole and maximal distribution, which were easily understood by the pupils.

- Fair distribution: each person should get the same number of objects.
- Whole distribution: only whole/complete objects should be distributed; otherwise, no objects should be distributed. That means that it is not allowed to break the objects and distribute the parts.
- Maximal distribution: each person should get the highest number of objects possible.

The use of a distribution approach, and notably these three key notions, were the basis of the proof-based teaching sequence to develop the mathematical content *Division and*

Divisibility. Both, division and divisibility, were defined in terms of the three key notions, so that the teachers and students could go back to them when needed. In short, we began with few initial fundamental notions and new assumptions, properties and concepts were built or derived from them. The knowledge construction became evident when students were capable of answering problems that demanded mathematical justifications of their answers. In the process of knowledge construction, it could be seen that students did not only participate actively, but were also encouraged to correct their classmates' or their own answers, refine ideas, suggest conjectures, and so on. All of this showed us that it was possible to develop a classroom environment rich in knowledge construction, in which the students experienced similar processes to those experienced by professional mathematicians, including especially the process of proving to discover and establish new knowledge. Proof-Based Teaching is a way of teaching mathematics that does not only has these goals, but also aims at bringing students' awareness about the mathematical activity involved and its nature in order to understand how and why mathematics works the way it does. Its major goal is to promote the development of independent thinkers, who can eventually construct their own mathematical knowledge without the support of their teachers.

In the two aforementioned research studies I combined three important features of Proof-Based Teaching (PfBT, see Reid & Vallejo-Vargas, 2017): (1) establishing a framework of established knowledge (the toolbox) from which to prove, where the "key notions" (as I called them in Vallejo-Vargas & Ordoñez-Montañez 2014) play a crucial role and are the basis of the proof-based teaching theories (see Reid & Vallejo-Vargas, 2019); (2) establishing an expectation that answers should be justified within this framework, which is a *sociomathematical norm* (Yackel & Cobb, 1996) that is set up within the classroom community; (3) proving is the means for constructing new mathematics and it necessarily involves the use of deductive reasoning.

It is also important to comment on the experience that I have preparing prospective elementary teachers at a private university in Peru, in the professional development of in-service (high and primary school) mathematics teachers, and training mathematics coaches (teachers who train other teachers). Thanks to these experiences, I have been able to observe the difficulties future and in-service teachers in Peru face while engaging in reasoning and proving, which are consistent with those pointed out by other researchers (see above). I have refined my teaching methods to address these difficulties over the years.

My prior research and experience have convinced me that it is crucial to start working on proof-based construction of mathematical knowledge from the beginning, which means with primary school students. However, there is still little in the preparation of primary school teachers that prepares them for PfBT, and in fact as I mentioned above, research has shown that engaging in proving and understanding the nature of proving and its role in mathematics and mathematics learning are weaknesses of future primary school teachers.

In this respect, the design, implementation and evaluation of research-based interventions that aim at supporting the teachers' development of proof-related skills is crucial. Although there have been some efforts to help *future* teachers to succeed in working reasoning and proving in their classrooms, not much has been explored in terms of the support *in-service* teachers need to successfully engage their students in proof-related activities. In addition, working with in-service teachers may have some advantages. For example, challenges that some prospective teachers faced while teaching proof in their mentor teachers classrooms included their lack of knowledge of students (see G. J.

Stylianides, et al., 2013). It involved the prospective teachers' lack of previous classroom experience and limited reflection in action, which made it harder for them to engage students in high-level mathematical activities. In-service teachers can anticipate some of the students' answers, difficulties, doubts, etc. Their prior generic teaching experience may alleviate part of the challenges that might emerge while teaching in a proof-based way.

There are many suggestions that have been made to support teachers' work; however, including all of them at once in a single intervention is not feasible and reasonable. In that sense, giving priority to some aspects and observe the effects of such a selection can be a path to follow. I am interested in researching an approach to teacher education that prepares primary school teachers to successfully engage in proof-based teaching. This approach is based on the work of A. J. Stylianides and Ball (2008), G. J. Stylianides and A. J. Stylianides (2009, 2014), Reid (2011), G. J. Stylianides, A. J. Stylianides and Shilling-Traina (2013) and my own prior research on relevant mathematical and pedagogical knowledge.

My research addressed the following research questions (RQs):

> RQ1: How do in-service primary school teachers' assumptions related to dis/proving change while engaged in an intervention focused on Proof-Based Teaching (PfBT) and understanding the nature of proving?
>
> RQ2: How are the in-service primary school teachers' assumptions that changed visible during their teaching in schools?
>
> RQ3: What design principles for a teacher development intervention focused on Proof-Based Teaching (PfBT) in primary schools can be abstracted from two cycles of such intervention?

In relation to RQ1, I additionally explore the question:

> RQ1a: What features of the intervention led to the observed changes?

While RQ1 and RQ2 have a descriptive character, RQ1a and RQ3 are more predictive.

Chapter 2: Literature Review

In order to put together the design of an intervention for teachers, I first needed to be aware of the specific challenges that learners in general, and (pre- and in- service) teachers in particular, face when engaged in proof-related activities. Likewise, I needed to be aware of previous research that had a focus on the design of interventions that aimed at fostering teachers' understandings so that they could engage themselves and/or their students in reasoning and proving.

I divide this chapter into three sections. Section I is devoted to the review of literature related to some of the challenges that may arise when engaged in proof-related activities. Section II has a focus on literature related to the knowledge for proof-related teaching and interventions designed where the development of part of that knowledge is promoted. Section III includes a review of the literature related to cognitive conflicts and proof.

I. Challenges when engaged in proof-related activities

In the context of the engagement in proof-related activities, several difficulties have been identified in different age groups. Here I consider eight major challenges: the converse error (Section 1); the lack of understanding about proving universal statements (Section 2); the reliance on the mode of expression when evaluating an argument (Section 3); the challenges involved in determining people's assumptions about proof (Section 4); the lack of understanding falsity and disproving of universal statements (Section 5); the lack of understanding about proving existential statements (Section 6); the lack of understanding falsity and disproving of existential statements (Section 7); and the lack of understanding about negation (Section 8). For every challenge and according to the existing literature I include possible explanations that might account for why they arise in the first place and, when possible, some directions for solutions.

1. The converse error

The *converse error* consists in inferring "*p*" given that "*p implies q*" and "*q*" (Epp, 2020, p. 73), or, in other terms, that based on the assumption that *"If A, then B"* (or *"All A are B"*) is true, an individual interprets it to mean that the converse (*"If B, then A"* or *"All B are A"*) is also true. It is also called *the fallacy of affirming the consequent* and is a common form of reasoning that individuals use (e.g., Anderson, 2015; Baggini & Fosl, 2010; Epp, 1999; 2003; Hoyles & Küchemann, 2002; Newstead & Griggs, 1983; Wason, 1966; 1968). The converse error is regarded as one persistent challenge students usually face when dealing with conditional statements (G. J. Stylianides & A. J. Stylianides, 2017). For example, Hoyles and Küchemann (2002) showed that over 60% of high-attaining year-8/year-9 school students interpreted conditional statements as their converses.

Particularly, this error has been shown when individuals of different ages and backgrounds solved the famous *Wason's selection task* (Wason, 1966), which has been widely used to test the ability to use deductive reasoning and, in particular, identify the cases that falsify a conditional statement. Valiña and Martín (2016) explained that the Wason's selection task "is a meta-inference problem which requires the understanding of a conditional 'if… then', and the formulation and verification of hypotheses" (p. 926). In essence, the original task is contextualized in such a way that its resolution is independent of practical knowledge (i.e., it is a content-free task). It has been shown that very low percentages of students managed to successfully solve the original task (e.g., Evans, 1972). Among the most common observations made by Wason (1966) about students' use of deductive forms of reasoning[1] were that most students used *modus ponens* (given *"if p then q"*, and *"p"*, then *"q"* is deduced) with confidence. In contrast, very few students managed to use *modus tollens* (given *"if p then q"*, and *"not-q"*, then *"not-p"* is deduced) and many students made the converse error and committed the fallacy of denying the antecedent (given *"if p then q"*, and *"not-p"*, then *"not-q"* is wrongly inferred).

Researchers can interpret converse errors differently. For example, in Buchbinder and Zaslavsky's (2011) study two top-level 10th grade students exhibited strong confidence in a (false) theorem that led to an incorrect proof. The two students seemed pretty confident that *any quadrilateral with perpendicular diagonals was a rhombus*. The authors claimed that the students confused it with a correct well-known theorem that *the diagonals of a rhombus are perpendicular* and explained that the students "manifested a distortion of the antecedent, done by applying the claim within distorted conditions… by omitting the condition that the diagonals need to bisect each other" (p. 279). However, an alternative interpretation of the students' conclusion could have been given in terms of the converse error. That is, the students' overconfidence possibly stemmed from their implicit assumption that, based on the true property of quadrilaterals *"If a quadrilateral is a rhombus, its diagonals are perpendicular"*, it followed that its converse *("If the diagonals of a quadrilateral are perpendicular, the quadrilateral is a rhombus")* was also true.

One direct effect of the converse error on individuals' proof-related activities is, for example, accepting a proof of the converse statement as if it were a proof of the original statement (e.g., Knuth, 2002a; Selden & Selden, 2003; Weber, 2010), or the individuals' disproof of the converse as if they were disproving the original statement (e.g., Zaslavsky & Ron, 1998).

Possible explanations

Possible explanations for the converse error involve either the logical structure of the statements, or the influence of context, practical knowledge and language.

Explanations involving the logical structure of the statements include a reading of "is a" as "is equal to", an assumption that all-statements are symmetrical, ignoring the implication entirely, reading "implies" as "and", relying on the truth value of the antecedent and the consequent, and the "atmosphere" of the statement.

One of the first explanations that emerged to make sense of the converse error is the ambiguity of the relational copula "is a", although Revlin and Leirer (1980) claimed that was empirically weak.

> Chapman and Chapman (1959) report that reasoners appear to encode "is a" as "is equal to" (an identity relation) rather than the logical "is included in" (proper inclusion) … One implication of this conception of conversion is that the universal affirmative, "All A are B," would be interpreted as "A equals B" (identity relation). (p. 447)

Instead, Revlin and Leirer (1980) suggested that a possible explanation could be that some individuals may assume that "all-statements" are symmetrical, just like "some-

[1] The study of the development of deductive reasoning has been of great attention in the field of psychological research on human reasoning (for a review, see Stylianides & Stylianides, 2008).

statements" and "no-statements" are. Revlin and Leirer claimed that "this notion of symmetry has become part and parcel of the explanations given to inference errors" (p. 448).

Bucci (1978) proposed the *structure-neutral interpretation* to suggest that individuals ignore the logical structure of a statement. That is, they tend to interpret a universal affirmative statement (UAS) as "a simple string or unordered set of substantive words without hierarchical structure" (p. 58); that is, a UAS of the form *"All X are Y"* is interpreted as *all, X, Y*, where the quantifier "all" is not applied to one of the terms (*X* or *Y*). According to Bucci, the use of this approach is age dependent (the younger the children, the more it is used) and it was manifested by the children's rejection of cases where one or both of the attributes (*X* and *Y*) were absent. Bucci also explored the effect of the content on the way statements were interpreted and found out that it played an important role and that it "can be accounted for on the basis of an underlying structure-neutral approach" (p. 59).

Durand-Guerrier et al. (2012) pointed out that students tended to confound implications ("*p* implies *q*") with conjunctions ("*p* and *q*"). In that case, that would explain why individuals see no difference between an implication and its converse since the order is irrelevant in conjunctions.

The truth value might also be a factor that explain individuals' decisions on the way they interpret "if-then" (or "all-") statements. Based on their longitudinal UK nationwide study on students' understanding of logical implication, Hoyles and Küchemann (2002) distinguished four categories of students' meanings for logical implication in terms of the relation with its converse, as well as the rationale behind. *Type A* responses suggested that both a logical implication and its converse were the same and did not use any particular data as a rationale; instead, these responses pointed to antecedent and consequent as interchangeable. The same truth value of both, the logical implication and its converse, (either both true or both false) justified in a way such conclusion. *Type B* responses also suggested that a logical implication and its converse were the same and also referred to the interchangeability of antecedent and consequent; however, in contrast to type A responses, they were characterized by their reliance on data that involved the truth values of antecedent and consequent, while they also referred to their interchangeability. Among those responses, some students concluded that the logical implication was false given that one condition (antecedent or consequent) was true and the other was false, and so the converse was also false. This type of response exhibited poor understanding of disproving implications. *Type C* responses are those that began as type A or type B responses and then when focused on the truth value of the implication and its converse (which were opposed), they shifted to conclude that a logical implication and its converse were not equivalent. *Type D* responses clearly stated that a logical implication and its converse were not the same. There is reference to the order ("the order matters") and to their differing truth values (while one is true, the other is false).

Another reason that can explain the converse error may be the universal or conditional atmosphere that the statements share. Woodworth and Sells (1935) suggested the notion of an "atmosphere effect". They explained that "[t]he atmosphere of the premises may be affirmative or negative, universal or particular. Whatever it is, according to the hypothesis, it creates a sense of validity for the corresponding conclusion" (p. 452). For example, a negative atmosphere in the premises may lead to make a negative conclusion, a universal atmosphere may account for universal conclusions, etc. According to the authors, the atmosphere influences the conclusions an individual makes. Notably, it could explain why some individuals accept that based on the universal statement *"All X are Y"* (or the conditional statement *"If X, then Y"*), one can conclude the universal statement *"All Y are X"* (*"If Y, then X"*).

The converse error could also be explained by language, context and practical knowledge (see Anderson, 2015; Valiña & Martín, 2016, for a state-of-the-art study on semantic and pragmatic factors that influence reasoning).

Baggini and Fosl (2010) claimed that "[the converse error] is a very simple mistake to make, since in everyday English, we distinguish between conditionals and biconditionals implicitly, by context, rather than by explicit stipulation" (p. 153). The context might explain some challenges. For example, it is interesting that when the original Wason's task was contextualized in a familiar setting (e.g., in a post-office context), where the individuals were asked to interpret the implication involved as a requirement, the results obtained were much better than those obtained with the original content-free task (e.g., Johnson-Laird, Legrenzi & Legrenzi, 1972, as cited in Durand-Guerrier et al., 2012; Cosmides, 1989). In the existing literature, this phenomenon has been called *the thematic facilitation effect* (see Evans, Newstead & Byrne, 1993, Chapter 4, for a review). Since then different factors have been modified in the original selection task, which led to different versions, in order to explain individuals' responses (see Valiña & Martín, 2016, for a review).

Another linguistic explanation relies on the interpretation of the verb "to be", which depends on the function it plays in the statement and may be implicit and therefore challenging when construing single-quantified (SQ-) statements. Among the distinctive linguistic features of language, Veel (1999) included the use of relational clauses in mathematics. For example, the verb *to be*, as in the examples "A number divisible by 6 *is* divisible by 3" and "A number divisible by 6 *is* a number that when divided by 6, leaves a remainder of 0", plays the role of a relational clause. According to Veel, there are two types of relational clauses: *attributive* and *identifying*. From the two examples I gave, in the former the verb *to be* functions as an *attributive* relational clause, while in the latter example it works as an *identifying* relational clause. While in the former example the meaning is that of inclusion (the set of numbers divisible by 6 is included in the set of numbers divisible by 3), in the latter example the meaning suggested is that of equality/identity, which is commonly used in definitions (the set of numbers divisible by 6 is equal to the set of numbers that when divided by 6, leave a remainder of 0). In contrast to the non-reversible relation suggested by the verb "to be" in the former example (the set of numbers divisible by 3 is not included in the set of numbers divisible by 6), the verb "to be" in the latter example is reversible. It is possible that some individuals confuse the role that the verb "to be" plays in these cases and the converse error emerges.

Practical knowledge influences the way adults reason about syllogisms. Higher numbers of correct answers for verbal syllogisms were presented when the conclusion and the participants' knowledge about the world were consistent. This interpretation is confirmed as the participants mostly provided incorrect answers to syllogisms where their practical knowledge fell short (e.g., Luria, 1976; Scribner, 1975). Epp (2003) speculated that a reason for why some people make the converse error is that

> they have come to take for granted that the truth of *if p then q* implies the truth of *if q then p* – unless their 'world knowledge' obviously contradicts this assumption, which is not usually the case when they try to analyze a new mathematical situation (p. 889)

Sometimes "if-then" statements in mathematics are construed as equivalent to their converses because of the way they are used in everyday conversations, where they also convey "if-not, then-not", as pointed out by Hoyles and Küchemann (2002). For example, *"If you clean your bedroom, you can go to the shopping mall with your friends"* is usually interpreted in common speech as *"If you do not clean your bedroom, you cannot go to the shopping mall with your friends"*. That is, in everyday speech conditional statements are usually interpreted as biconditional statements (Epp, 2003). However, in mathematics a conditional statement (*"If X, then Y"*) is not necessarily equivalent to its converse (*"If Y, then X"*). Similarly, expressions such as *"It is necessary to do A to get B"*, according to Durand-Guerrier et al. (2012), are interpreted in everyday language like *"If I don't do A, then I won't get B"* and *"If I do A, then I will get B"*; however, in mathematics only the former interpretation is valid, not the latter.

The way tasks are organized in research studies might also explain the forms of reasoning used. Hawkins, Pea, Glick and Scribner (1984) investigated the effects of problem complexity, problem content, and task organization on preschoolers' (4- and 5-year-old children) performances when solving syllogistic problems. They showed that the task organization as well as the empirical truth value of the problem content played a crucial role in the kids' use of their deductive reasoning skills. When *fantasy problems*[2] were presented in an initial position, the kids tended to respond more logically to these problems because practical knowledge did not interfere with their performance. When either incongruent or congruent (with practical knowledge) problems were presented first, they appeared to influence the way kids solved the subsequent tasks, which tended to show an empirical bias.

A lack of instruction on in/valid inferences might also explain low-achieving performances in reasoning tasks. Deductive reasoning rules like *modus ponens* seem to be, if not intuitive, natural to adopt. Nevertheless, other deductive modes of reasoning involved in dis/proving demand work and time to be accommodated and assimilated. Durand-Guerrier et al. (2012) stated that the only rule of inference that is explicitly taught in France is *modus ponens* in grade 8 (13-14-year-old students). The authors pointed out that this kind of curricula is partly the reason "why students arriving at university both in France and elsewhere have severe problems with the logic needed for learning more advanced mathematics, especially the logic that involves quantification" (p. 373).

Possible solutions

The problem of the converse is one of the widespread weaknesses for which G. J. Stylianides and A. J. Stylianides (2017) called for research that aims at the design of classroom-based interventions and noted that no meaningful advances in this area had yet been taken in this area. Nevertheless, some hints for what to consider in future interventions have been given by some researchers. Epp (2003) and Durand-Guerrier et al. (2012) suggested exploiting similarities between formal and everyday language. They pointed out that even though both have many differences, there may be cases where the everyday and mathematical interpretation of statements coincide. To illustrate this, Epp provided the example *"If Sam lives in Chicago, then Sam lives in the U.S."* and highlighted that it does not necessarily imply its converse, *"If Sam lives in the U.S., then Sam lives in Chicago"*. Epp claimed that using such examples of statements, where their mathematical interpretation is in accord with its everyday interpretation, might support the students' enhancement of their mathematical reasoning skills. Durand-Guerrier et al. (2012) drew attention to the example of a true mathematical statement, *"For every real number x,* $x > 1$ *implies* $x^2 > x$*"*, for which its converse is not true (think about $x =$

[2] Hawkins et al. (1984) used the expression *fantasy problems* to refer to those problems "in which premises described mythical creatures foreign to practical knowledge" (p. 586).

-5); however, if the same true conditional statement is defined in the set of natural numbers, it actually becomes a true bi-conditional statement.

The use of Venn/Euler diagrams[3] to represent and analyze universal conditional statements may support overcoming the problem of the converse. Deloustal-Jorrand (2007) highlighted three points of view in which implication can be understood: sets, formal logic and deductive reasoning. She suggested that a lack of awareness about the linkage among them might explain most of the difficulties and mistakes when using implications. In Newstead and Griggs' (1983) study, approximately 75% of the undergraduate students chose the two correct diagrams that could represent the statement *"All As are Bs"*; namely, A is a subset of B and A is identical to B. This result was quite surprising to Newstead and Griggs because of the converse problem, as they expected that the students only chose the diagram where A was identical to B.

The converse error is directly related to the interpretation of universal/conditional statements. Next, I focus on some issues that emerge when proving universal statements.

2. Lack of understanding about proving universal statements

Two common challenges that might arise when engaged in proving universal statements (USs) are the acceptance and production of empirical arguments to prove USs and the belief that proof is a fallible construct. Both reflect a lack of understanding of proving USs.

The value of generating examples in mathematics has been underscored by different researchers (e.g., Ellis, Lockwood, Dogan, Williams & Knuth, 2013; Lockwood, Ellis & Lynch, 2016; Morselli, 2006; Watson & Mason, 2005; Zazkis & Chernoff, 2008; Zazkis, Liljedahl & Chernoff, 2008). For example, the use of examples is relevant in problem solving as suggested by mathematicians (Polya, 1945); requesting students to generate their own examples can be beneficial when they learn about new concepts (Dahlberg & Housman, 1997); examples play an important role in making conjectures (Zazkis et al., 2008); examples can support the creation and resolution of cognitive conflicts (Zazkis & Chernoff, 2008).

In contrast, some studies that focused on the effect that example-based reasoning can have in proving, showed no significant benefits. Iannone, Inglis, Mejía-Ramos, Simpson and Weber (2011) observed that generating examples of mathematical concepts did not improve undergraduate students' performance when proving conjectures related to those concepts. The authors did not find a "significant difference between the proof production success of students who generated examples compared to those who studied worked examples" (p. 9).

The relation between the use of examples and proof has been widely investigated. One solid finding in the argumentation and proof literature is that most individuals (with different ages and backgrounds) tend to use examples when asked to prove infinite universal statements; however, the reasons they do so may differ. Research has shown that individuals tend to not recognize that empirical (example-based) arguments in mathematics are inconclusive and accept them as valid proofs (e.g., Chazan, 1993; Harel & Sowder, 1998; Healy & Hoyles, 2000; Martin & Harel, 1989; Morris, 2002; A. J. Stylianides & G. J. Stylianides, 2009).

[3] See Chapter 3, Section II.2 for details related to the representation of SQ-statements with Venn/Euler diagrams.

In Martin and Harel's (1989) study, preservice elementary teachers were requested to rate both inductive (example-based) and deductive arguments for their validity. They found out that many accepted inductive arguments as valid proofs. Furthermore, as the authors pointed out, "over 46% simultaneously rated general proof and at least one of the inductive-argument verification types high" (p. 48), showing that those preservice teachers did not find those frames mutually exclusive.

Morris (2002) observed that Martin and Harel (1989) had not established whether the teachers who accepted both inductive and deductive arguments as valid proofs could distinguish the two types of arguments. Morris' study showed that only 30% of preservice elementary and middle school teachers in her study distinguished between inductive and deductive arguments and they "experienced deductively derived conclusions as necessary and inductively derived conclusions as uncertain" (p. 110). 40% could not distinguish the two types of arguments. Those preservice teachers considered that both types of conclusions were necessary. The remaining 30% of the preservice teachers made the distinction between inductive and deductive arguments but regarded both inductive and deductive conclusions as uncertain or indeterminate.

There might be cases where individuals reject example-based arguments; however, the reasons for why they do so may also reflect a lack of understanding of proving USs. They might base their reasoning on an authority such as a teacher who taught them to discard all example-based arguments when proving USs. For example, Tabach, Levenson, Barkai, Tirosh, Tsamir and Dreyfus (2010b) found that all in-service secondary school teachers who participated in their study produced correct symbolic proofs for the true given universal statement *"The sum of any 5 consecutive natural numbers is divisible by 5"*. Additionally, when judging numerical justifications, they all rejected that one supporting example could prove a universal statement. Although the authors did not explore the participants' form of reasoning to determine the sources and their nature, it is possible that at least some of those teachers learned that examples do not prove as a rule.

In Barkai, Tsamir, Tirosh and Dreyfus' (2002) study 27 elementary school teachers were given statements about divisibility that involved infinite cases (three universal, three existential, some true and some false). The teachers were asked to determine their truth value, justify their answers and decide whether their lecturers would accept such justifications as proofs. Barkai et al. showed that 52% of the teachers provided confirming examples to justify a true universal statement, from which 33% perceived that their justifications were proofs. On the other hand, 41% of the teachers gave correct deductive justifications, from which 33% were algebraic justifications and 30% of them regarded them as proofs.

Healy and Hoyles (2000) investigated the conceptions of proof in the context of arithmetic/algebra held by high-attaining 14-15-year-old students. In their proofs, the students mostly used examples; however, they were aware that those arguments would not receive the highest mark from their teachers, which suggested that the students were aware of the limitations of empirical arguments. More than 50% knew that empirical arguments only proved the statement for a subset of all cases involved in it.

In the same vein, Weber (2010) presented similar findings from his study with mathematics majors. In general, Weber found that the majority of the participants did not find the (only one) given empirical argument as convincing, neither did they consider it met the standards of a proof.

Another challenge when engaged in proving USs is the belief that a proof is a fallible construct. In this regard, there is evidence that many individuals do not acknowledge the conclusiveness of deductive arguments in mathematics (e.g., Chazan, 1993; Fischbein & Kedem, 1982; Knuth, 2002a; Morris, 2002; Vinner, 1983).

Possible explanations

The everyday context way to “prove” a conjecture may influence the way a conjecture is proved in mathematics. In that respect, Harel and Sowder (1998) pointed out that “[s]ince people’s evaluation of hypotheses in everyday life is probabilistic in nature …, the use of inductive evidence is only natural” (p. 252).

It is also possible that teachers are contributing to a reinforcement of an empirical proof scheme. Harel and Sowder (1998) claimed that the prevalence of the empirical proof scheme among students might be explained in terms of their teachers’ usage of such a proof scheme when teaching mathematics. In that case, students do not question their teachers’ modes of argumentation because in a way the teachers represent the authority in the classroom.

A lack of awareness of valid methods of proof may be also an explanation for why empirical arguments are so commonly used when proving USs. Healy and Hoyles (2000) found that students can be aware of the limitations of empirical arguments, but still they may not manage to provide anything better. Harel and Sowder (1998) claimed that students may lack knowledge about valid proof schemes and so they resort to what is accessible to them. That often leads them to rely on empirical proof schemes.

Besides that, it has been observed that typically students do not have strategic knowledge to prove (e.g., Weber, 2001). For example, Weber (2001) found that mathematics doctoral students had strategic knowledge that undergraduate students did not have. Weber pointed three types of strategic knowledge that the doctoral students seemed to use when proving: knowledge of the domain’s proof techniques, knowledge of which theorems are important and when they will be useful, and knowledge of when and when not to use syntactic strategies. According to Weber, the undergraduate students appeared to lack such strategic knowledge, which was exhibited as the several difficulties they faced while they engaged in proving.

The individuals’ introduction of personal knowledge might mislead them to judge empirical given arguments as proofs. Morris (2007) explored the factors that affected preservice teachers’ evaluation of mathematical arguments. Among other factors, she found that “[m]any preservice teachers did not limit themselves to what was written/said in a mathematical argument” (p. 508). Those future teachers did not succeed in monitoring the introduction of their personal knowledge when they evaluated the validity of given arguments. In many cases, the prospective teachers’ awareness of the truth of the statement acted as an intruder for their evaluation of the argument, as it was pointed out in Morris (2002).

In contrast, the fact that some individuals reject empirical arguments as proof may suggest previous explicit instruction in this particular issue. In some cases, students may be taught that proving an infinite universal statement through the verification of a few of the involved cases is not sufficient. In that case, students learn in an authoritarian way to reject empirical arguments. For example, even though it is not expressly stated in Barkai et al.’s (2002) study, the elementary school participating teachers seem to have had a prior explicit instruction on this topic and, notably, on what qualifies as proof and what does not. My presumption is based on the features of most of the arguments they provided, which were algebraic, even for the case of existential statements as well as their common assumption that examples would be rejected as proof by their instructors. Whether these teachers learned that empirical arguments were not proof in an authoritarian way or not, their assumptions about the nature of mathematical dis/proof is a factor that influences their rejection or acceptance of empirical arguments. For instance, many of the preservice teachers in Morris' (2002) study believed that an exhaustive verification of all cases involved in an already proven statement is the only way to establish the truth of the statement.

Weber (2010) sketched several hypotheses for why the mathematics majors in his study apparently did not to hold an empirical proof scheme. One had to do with the types of tasks that participants had to solve. The tasks only requested participants to evaluate arguments and the author inferred proof schemes from those evaluations. He pointed out that other proof schemes may have been inferred if the participants were engaged in, for example, constructing arguments. A second factor was the perception of generality in the empirical argument. In Weber's study the participants do not seem to have perceived a given empirical argument as general. That might have supported the participants' decision on rejecting the empirical argument. Weber also emphasized the students' previous engagement in a transition-to-proof course, where they might have learned to be cautious about empirical arguments. The author suggested that it might be the case that those who ended up participating in his study were those who did not hold an empirical proof scheme.

Individuals may believe that once a statement was proved to be true, it is possible that counterexamples could be found (e.g., Chazan, 1993; Knuth, 2002a). Knuth (2002a) explored secondary school mathematics in-service teachers' conceptions of proof. Most of the participants in his study correctly identified valid proofs; however, many of them did not discard the possibility "that counterexamples or other contradictory evidence may exist", which suggested their belief that "a proof is a fallible construct" (p. 401).

Another aspect of proving, where challenges might arise, is related to the evaluation of arguments. In terms of the evaluation of arguments and whether they qualify as proofs or not may, it be influenced by the mode of expression used in the argument. In the following section I focus on this issue.

Possible solutions

Research with mathematicians and their process of generating examples when proving have hinted ways that can support students to make more effective use of examples in their proof activity. Lockwood et al. (2016) investigated mathematicians' example-related activity when engaged in exploring and proving mathematical conjectures. Lockwood and her colleagues reported different criteria the mathematicians used to choose examples and the role those examples played during their proving process. In particular, the authors highlighted three characteristics that were crucial for the mathematicians' successful use of examples: the mathematicians were aware of their choices and the use of examples; the mathematicians understood the nature of the back and forth relationship between proof and disproof; the mathematicians were persistent and self-confident in their work. In addition, the authors pointed out that,

> [m]uch of the current literature on teaching proof in school mathematics underscores the goal of helping students understand the limits of such example-based reasoning… and typically characterizes example-based reasoning strategies as obstacles to overcome. (p. 166)

In fact, research shows that certain forms of example-based reasoning can be used when proving. For example, *generic examples* (see Mason & Pimm, 1984) can be used in arguments that qualify as proofs and so they can serve as a way to move from empirical forms of reasoning towards deductive ones. Some researchers have suggested that using generic arguments that qualify as proofs could be an alternative way to introduce novice to proof (e.g., Kempen, 2018; Reid & Vallejo, 2018; G. J. Stylianides et al., 2017). G. J. Stylianides et al. (2017) gave two reasons for why generic proofs are important: (1) they can serve as a bridge between empirical arguments and nonexample-based deductive arguments; and (2) generic arguments might be more accessible to students than conventional proof, because of their specificity and concreteness. Nevertheless, whether a generic argument qualifies or not as a proof may be a matter of perspective (Reid & Vallejo, 2018) and attention should be paid to this aspect. For example, Dreyfus (2000) expected that the teachers who participated in his study assigned a better grade to what *he* considered a generic proof; however, most of the teachers made negative comments about it and gave it low grades. Only about a third of the teachers could appreciate the generality of the argument. Reid and Vallejo (2018) pointed out that for a reader (e.g., a teacher) an argument that involves the use of some examples might have potential to become a (generic) proof; however, whether the author (e.g., a student) of the argument "sees" it in the same way, and intended to present the argument as a generic proof, may be uncertain without directly asking the author. Reid and Vallejo suggested that authors of generic arguments should include two types of evidence in written work to show that their attention is on the general aspects of the examples they use; namely: evidence of awareness of generality and mathematical evidence of reasoning.

In terms of interventions designed with the aim to support future teachers' emergent realization of the limitation of empirical arguments as valid methods to prove mathematical generalizations, G. J. Stylianides and A. J. Stylianides (2009) made some progress. They designed and implemented an intervention for pre-service teachers. The authors highlighted four features of the theoretical framework that underpinned the design of their instructional sequence: (1) proof as an integral element of engagement with mathematics as a sense-making activity; (2) inextricable relationship between instructional sequences and learning trajectories; (3) cognitive conflict as a mechanism for supporting developmental progressions in students' knowledge; (4) means for supporting the resolution of cognitive conflict and the role of the instructor. The sequence the authors designed entailed provoking cognitive conflicts for the prospective teachers to transit from an empirical argument justification scheme to a conventional/mathematical one, by going through a crucial experiment justification scheme (Balacheff, 1988). The authors stressed important aspects in this process. The importance of a conflicting case to create a cognitive conflict involved that it agreed with the pre-service teachers' example spaces for validation and that they became aware of their existing conceptions along this process.

3. Preference for a formal mode of expression when evaluating an argument

Considering the different modes of argument representation or expression (e.g., verbal written/oral, symbolic/algebraic, pictorial) in which proof may be presented, some studies have shown that individuals tend to reject acceptable non-formal modes of expression and some base their acceptance merely on a symbolic/algebraic representation (e.g., Barkai et al., 2002; Dreyfus, 2000; Knuth, 2002a; Tabach et al., 2009, 2010a).

Dreyfus (2000) reported on in-service high school teachers' views of proof. The teachers were given nine arguments for the universal statement *"The sum of any two even numbers is even"* and were requested to grade them and justify the grades they assigned. Several teachers criticized a *verbal* but complete proof because of the lack of formal language involved in it. That means that their judgement was based on their assumption that mathematical proofs should be in mathematical symbolic language. One of the teachers recognized the non-mathematical language used in the argument; however, to this teacher the use of formal language did not seem to be a requirement for a proof.

Knuth (2002a) found that a group of secondary school teachers who accepted the proof of a converse as if it were a proof for the original statement "seemed to focus solely on the correctness of the algebraic manipulations rather than on the mathematical validity of the justification" (p. 392).

Tabach et al. (2010a) investigated high school in-service teachers' use, acceptance and their expectations for students' use of *verbal* justifications. First, none of the teachers used a verbal mode of argument to justify the six (3 universal and 3 existential elementary number theory) given statements. Half or fewer used a single counterexample to refute the two false universal statements, whereas the other teachers gave symbolic general proofs. All the teachers relied on a symbolic mode of argument representation for the true universal and the false existential statements. Second, half of the teachers rejected valid verbal justifications. Moreover, some teachers had difficulties understanding the correct verbal justifications that were given to evaluate, but they did not reject them in principle. Third, most of the teachers did not expect a verbal mode of argument in their students' justifications. In this respect, Tabach and his colleagues stated that their findings "contrast with what is known from the literature about high school students' preference for a verbal representation for justifications due to their explanatory power." (p. 1088). Nonetheless, teachers may not be aware of their students' preferences for the modes of expression of their justifications. Indeed, Healy and Hoyles (2000) found that 14-15-year-old high-attaining students preferred *verbal* over symbolic justifications when proving; however, students perceived that those would get lower grades from their teachers and, instead, symbolic/algebraic justifications would get the highest grades. Nevertheless, the students did not choose algebraic arguments as one of their own approaches to prove. Students found algebraic arguments difficult to follow and so they could hardly decide whether an algebraic argument was valid. On the other hand, narrative arguments were more popular in the students' attempts to proof constructions and students were more successful constructing proofs when they used a narrative form. The students found it easier to evaluate narrative arguments. Only 18% rejected their explanatory power.

Possible explanation

It is possible that, as in the case of empirical arguments, an authoritarian instruction fosters a conception that proofs can be only presented in a formal mode of expression.

Identifying individuals' understanding of proving USs through their production and evaluation of arguments is useful; however, there might be some other aspects that can be involved and therefore may make it difficult to find out the assumptions that individuals use about proof. In the following section I focus on this issue.

4. Challenges involved in determining individuals' assumptions about proof

When determining individuals' assumptions about proof, and in particular their proof schemes (e.g., Balacheff, 1988; Harel & Sowder, 1998, 2007), analyzing the justifications they produce has been a common approach used by researchers. Nonetheless, to better understand students' assumptions about proof, that is not enough. Hoyles and Küchemann

(2002) claimed that the meaning students assign to the justifications they provide may not be as clear as one might think. As they put it,

> [t]his calls into question what students mean when they present even a correct counter example. Is it in fact showing a recognition that the statement is not true or that the statement is *sometimes* not true? (p. 218)

Furthermore, interpreting individuals' construction of arguments that fall short as proofs as their unawareness of the limitations of their arguments may be misleading, given that they may not know better approaches (e.g., Healy & Hoyles, 2000; Knuth, Choppin & Bieda, 2011; A. J. Stylianides & G. J. Stylianides, 2009). As pointed out by Knuth et al. (2011),

> [s]uch results not only suggest that students may not actually believe that examples constitute proof, but also that there is a difference between students' proof production competencies and their proof comprehension competencies. (p. 168)

There are cases when individuals consider that given empirical arguments qualify as proofs; however, the reasons for why they might do so cannot be simply reduced to assuming that they have an empirical scheme of justification. This is the case of some of the students who participated in Healy and Hoyles' (2000) study. Many of the students who incorrectly evaluated empirical arguments as valid proofs, did so because they were already convinced of the truth of the statement.

In her study with preservice teachers, Morris (2007) showed that certain classroom conditions may affect the assumptions about proof they use. Morris separated the prospective teachers into two groups: the V (valid) group and the NV (not valid) group. In contrast to the NV group, the V group received a general deductive argument among the list of arguments they were given to evaluate. The preservice teachers from the V group tended to use mostly the "explain why" criterion. It consisted in concluding that an argument that included a key idea through the use of one example was a valid proof, while an example-based inductive argument was not. On the other hand, the NV group used the "prove" criterion. It consisted in concluding that example-based arguments that included several confirming examples were valid and key-idea arguments were not since it only included the testing of a single example. Morris observed that very few of the preservice teachers manifested that they were looking for general valid arguments.

Individuals' evaluation of arguments is directly influenced by their previous proof-related learning experiences (e.g., Weber, 2010) and their assumption that a convincing argument and a mathematical proof may not be the same (Segal, 1999). Weber (2010) showed that many of the mathematics majors who participated in his study found that the two given arguments that relied on diagrams were convincing and could constitute proofs; however, many of them regarded them as non-rigorous proofs. From those who did not think that diagrammatic arguments were rigorous proofs, several pointed out "the presence of the graph as the reason for their judgement" (p. 323), because they learned that graphs were not allowed to be used in proofs.

Possible explanations

The fact that there might be many factors to be considered when determining individuals' assumptions about proof can explain why it is such a big challenge. For example, individuals may find it easier to evaluate arguments rather than construct them on their own (e.g., Healy & Hoyles, 2000; Reiss, Hellmich & Reiss, 2002). Healy and Hoyles (2000) concluded that "students were significantly better at choosing correct mathematical proofs than at constructing them" (p. 407). Nevertheless, as Morris (2007) have pointed out, individuals may introduce their personal knowledge about the truth value of a statement when evaluating arguments, sometimes without them even realizing it.

It may be that students do not understand the proofs they are given to evaluate and that makes it difficult to access their assumptions about proof. For example, Mejía-Ramos et al. (2012) developed a model to assess undergraduate mathematics students' proof comprehension, which they built on Yang and Lin's (2008) work. Mejía-Ramos and his colleagues considered two levels of proof comprehension: *local* and *holistic*. The former addressed students' understanding of one or a small number of statements included in the proof. It included three types of assessment: meaning of terms and statements, logical status of statements and proof framework, and justification of claims. For the latter, their model consisted of four types of assessment: summarizing via high-level ideas, identifying the modular structure, transferring the general ideas or methods to another context, and illustrating with examples. G. J. Stylianides et al. (2017) pointed out that "[r]esearch into how students *understand* proof is limited" (p. 243).

Possible solutions

A. J. Stylianides (2019) suggested the need to use different research methods for assessing students' mathematical abilities such as that of proof construction; for example, consider students' proof constructions in different modes of argument representation (e.g., written and oral). A. J. Stylianides investigated the role of the *written* versus the *oral* mode of argument representation in the proof production of secondary school students[4]. A. J. Stylianides found out that "all strong written arguments retained their status during their oral presentations, while all weak written arguments were upgraded to strong arguments during their oral presentations" (p. 171).

Having access to the students' perception on whether they regard their own arguments as proof or not and why is also crucial. This is precisely one aspect that G. J. Stylianides and A. J. Stylianides (2020) underscored that some of prior research on this topic seems to have overlooked when identifying individuals' justification schemes; namely: *Individuals' own perceptions of whether their produced arguments actually met the standards of proof.*

Proving is closely related to disproving. Hence, becoming aware of the challenges that might emerge when individuals engage in disproving is also a crucial point to consider. In the following section I include some of those issues.

5. Lack of understanding about falsity and disproving of USs

In this section I focus on three common challenges that are reported in the existing literature in relation to the understanding about falsity and disproving of universal statements. One is the individuals' lack of understanding of what is the sufficient justification to disprove a US (Section 5.1). Another challenge is the individuals' non-acceptance of more than the minimal disproof for USs that is a counterexample (Section 5.2) and, finally, the difficulties individuals may experience to find counterexamples for USs (Section 5.3). The three challenges reveal individuals' limitations in their understanding of false USs and what is involved in disproving them.

[4] Before the students engaged in constructing proofs, they were introduced to the concept of proof and discussed what counted as a proof.

5.1. Lack of understanding about the sufficiency of one counterexample to disprove USs

From a mathematical point of view, one counterexample is sufficient evidence to disprove a US (see Chapter 2, Section 2.3). Research has evidenced that some individuals do not seem to be convinced of or do not understand the sufficiency of one counterexample to refute a universal statement.

Less than half of the in-service secondary school teachers who participated in Tabach, et al.'s (2010b) study produced one counterexample for the two false (divisibility) universal statements they were given. The other teachers produced symbolic justifications. When the teachers were given numerical justifications to be evaluated, they all correctly identified that one counterexample was sufficient to refute those two false USs. On the other hand, half of the elementary school teachers who used counterexamples to disprove that *"The sum of any four consecutive integers is divisible by four"* in Barkai et al.'s (2002) study regarded them as proofs, whereas the other half doubted the sufficiency of those examples to refute the statement. Some of them claimed that examples were not mathematical proofs, but since they did not know what the proof was, they provided some examples. Similarly, some in-service secondary school teachers in Giannakoulias, et al.'s (2010) study only used counterexamples to complement a theorem they referred to in their disproving or when they could not use a suitable theorem.

In the context of school students, Zaslavsky and Ron (1998) found out that only 10% of the high achieving high school students in their study correctly used counterexamples in more than one case and two thirds "either did not find it appropriate to use counterexamples or were not able to correctly use counterexamples for any statement" (p. 230).

Balacheff (1991) identified six ways in which 13-14-year-old students treated refutations: rejection of the conjecture, modification of the conjecture, the counterexamples considered as an exception, introduction of a condition, definitions revisited, and rejection of the counterexample. Vinsonhaler and Lynch (2020) reported on middle and high school students' similar behavior. Some students presumed that a universal-statement conjecture was true, but when confronted with a counterexample they modified the conjecture so as to exclude the counterexample, which the authors noted that was similar to *monster barring*, a phenomenon that Lakatos (1976) had described.

Possible explanations

As in other of the previous cases, authoritarian instruction might be one of the explanations for why individuals reject that one counterexample is sufficient to disprove a US. Concretely, students might hold this assumption because they were taught that "examples" do not prove, as might have been the case in Barkai et al.'s (2002) study. Furthermore, in some cases individuals possibly learned that mathematical theorems are only reached by means of other theorems and not examples, as in Giannakoulias, et al.'s (2010) study, who claimed that some teachers seemed to have presumed that "refutation by using theorems provides stronger and more general conclusions than by using counterexamples and that counterexamples are exceptions in the sense that Lakatos (1976) discusses them" (p. 166). That is precisely another explanation for why some individuals do not acknowledge conclusiveness to the case of counterexamples when refuting USs. Counterexamples are sometimes seen as exceptions or special cases and therefore insufficient to disprove a universal statement. Reid (2002) reported on fifth graders' experiences with counterexamples and how they often considered them as exceptions for a statement given, suggesting that counterexamples are not sufficient to disprove the statement.

Another factor that can explain why individuals struggle to understand the sufficiency of one counterexample to refute a US may be related to the truth value of the statement, especially when it admits both confirming and contradicting examples. Individuals might be unsure of the truth value for such statements. For example, Barkai et al. (2002) showed that when given a false universal statement that admitted confirming examples (*"The sum of any three consecutive integers is divisible by three"*), some elementary school teachers (8%) struggled to determine whether it was true or false because they could find both supporting examples and counterexamples. Other teachers (23%) claimed that it was true, while the others (69%) recognized that it was false. From the first group, none of the teachers expected that their justifications would be accepted as proofs. From the second group, from those who provided supporting examples, only one teacher expected that his/her justification would be accepted as a proof. Most of the third group (50%) provided one or more counterexamples, most of which considered that they would be considered as proof; however, some other teachers were unsure because they claimed that examples were not proofs. The rest of the third group (19%) provided algebraic proofs. Vinsonhaler and Lynch (2020) found out that for some students the number of (confirming or counter) examples matters, which leads them to conclude that a (false) conjecture is true if more confirming examples are presented, or that it is false if more counterexamples are shown. This kind of behavior clearly reveals individuals' lack of understanding of the nature of counterexamples. Buchbinder and Zaslavsky (2009) explored the status that one tenth-grade student attributed to examples when determining the validity of statements. The fact that the given universal statements allowed both confirming as well as contradicting examples puzzled the student, which at some point misled her to concluding that the statement was both true and false.

There is evidence that sometimes individuals are persuaded that a generalization is false only with the support of *certain* counterexamples. Zazkis and Chernoff (2008) described the misconceptions of two prospective elementary school teachers and the role that pivotal and bridging examples played in their process of overcoming their misconceptions. In concrete, Selina, one of the prospective teachers, assumed that 437 was prime, despite being aware that $437 = 19 \times 23$. To support her incorrect assumption, Selina used her (wrong) assumption that *the product of two prime numbers was a prime number*. Still, when the number 15 was suggested, Selina recognized that it was not prime, but rejected it as a counterexample for her assumption. She explained that it was different from 437 since 15 involved the multiplication of "building blocks" prime numbers (3 and 5), unlike 437 that involved numbers 19 and 23. Nonetheless, the number 77 did count as a counterexample for Selina. It played the role of a bridging example for the prospective teacher as it "is 'small enough' that it is similar to 6 and 15 (but unlike 437) because its factors are immediately recognizable… and it is not composed of 2, 3 or 5, the numbers that Selina referred to as 'building blocks'" (p. 201).

In the same vein, Vinsonhaler and Lynch (2020) provided evidence that even though some middle and high school students were aware that one counterexample was sufficient to disprove a US, they regarded certain counterexamples as "better" than others.

5.2. Rejection of non-minimal disproofs

From a mathematical point of view, one counterexample is sufficient/minimal justification to disprove a universal statement; however, non-minimal disproofs also count as valid disproofs (see Chapter 2, Section 2.3). The acceptance of non-minimal disproofs as valid, though, seems to be questioned by some individuals. About 25% of the 50 in-service secondary teachers who participated in Tabach et al.'s (2010b) study rejected non-minimal disproofs in the form of two or more counterexamples to refute two false universal statements; the others accepted them. Tsamir et al. (2009) reported the case of one secondary-school teacher's dis/proving of and evaluation of the correctness of 43 justifications for six (universal/existential, true/false) statements. The justifications were presented in different forms of expression (numerical, verbal and symbolic) and different modes of argumentation (one counterexample, several counterexamples, general argument, one supporting example, etc.). The teacher who participated in the study rejected all correct arguments for the case of false universal statements that included more than the "minimal" justification.

In contrast, there might be cases where students provide or accept a correct non-minimal disproof for a generalization. For example, in Tabach et al.'s (2010b) study around half of the teachers produced a non-minimal symbolic justification to refute the given universal statements.

Possible explanations

Students may reject non-minimal disproofs for USs since they may not conform with patterns of disproof they learned. Tsamir et al. (2009) summarized the teacher's rationale to behave the way she did in the following terms: "if the justification does not follow the needed framework, it may reflect a lack of understanding by the author of the justification, and hence she rejects it." (p. 64).

For those cases of students who provide/accept a correct non-minimal disproof for a generalization, a possible explanation may also reflect their lack of understanding of when a universal statement is false. Epp (1999) referred to one of her previous studies in which she found out that a large sample of students selected the statement *"No mathematicians wear glasses"* as the statement that conveyed what it meant for the statement *"All mathematicians wear glasses"* to be false. This means that some individuals may assume that in order to disprove a universal statement (e.g., *"All X are Y"*), they need to prove its contrary statement (i.e., *"No X is Y"*). Notably, Barkai et al. (2002) showed that some elementary school teachers (algebraically) disproved that *"The sum of any four consecutive integers is divisible by four"* by proving that *the sum of any four consecutive numbers was not divisible by 4*. This means that they proved its contrary statement *"No sum of any four consecutive integers is divisible by four"*.

It is also possible that students who provide non-minimal disproofs are actually aiming at more sophisticated disproofs. For example, Lew and Zazkis (2019) explained that there may be two reasons that can explain why undergraduates generate evidence that goes beyond the use of one counterexample when disproving USs: to generate a more general counterexample and to write a simpler cleaner proof.

5.3. Difficulties finding counterexamples

Individuals may have difficulties finding counterexamples for false universal statements. For example, Zaslavsky and Peled (1996) found out that only one third of the participant in-service secondary mathematics teachers could provide at least one complete, well-justified, and correct counterexample for the false universal statement *"Any commutative operation is also associative"*. Furthermore, only 4% of the pre-service teachers in the study could provide a counterexample.

In some cases, because of the struggles to find a counterexample, individuals attempt instead to prove a false universal statement. Ko and Knuth (2009a) found out that while some of the undergraduate majors who participated in their study could provide an adequate counterexample for two false universal statements, others tried to prove that those statements were true, and a considerable number of participants did not provide any response. In Potari, Zachariades and Zaslavsky's (2010) study only about 17% of the teachers correctly rejected a false statement about congruent triangles and some teachers considered that the statement was true. Among those who realized that the statement was false, there were teachers who either resorted to known theorems, or counterexamples. In the former case, few teachers provided a valid proof and the others "believed that the non-applicability of the known relevant theorems implied that the claim was wrong" (p. 287). In the latter case, some teachers only pointed out the need to provide a counterexample; however, they did not provide any. From those teachers who tried to construct a counterexample, most of them did not manage to provide a correct counterexample.

Possible explanations

A lack of awareness of the characteristics that counterexamples should have may explain why it is difficult for some individuals find counterexamples. The importance of being able to construct a description of all possible counterexamples has been emphasized, for instance when indirect reasoning is involved (e.g., Yopp 2017). There is evidence that some students do not seem to be aware of the conditions that counterexamples should satisfy. For example, Zaslavsky and Ron (1998) reported that many top-level high school students who accepted a counterexample as sufficient evidence to disprove a universal statement accepted/used supporting examples or counterexamples for the converse statement as if they were counterexamples for the original statement. In their study, Buchbinder and Zaslavsky (2007) included the true statement *"The domains of function $f(x)$ and its derivative $f'(x)$ are not necessarily the same"*. Some high school students considered that it was a false statement and pointed to the case $y = \sqrt{x}$ to support their answer, even though it is a confirming example, not a counterexample. Ko and Knuth's (2009b) study focused on the dis/proving of prospective secondary school teachers. Two of the future teachers attempted to disprove a false US; however, the example they chose did not satisfy the conditions for it to be a counterexample. Six of them believed that a (true) statement was false and tried to provide counterexamples for it. Again, they did not exhibit an understanding of the conditions a counterexample should have. These prospective teachers revealed not only a poor understanding of dis/proof, but also of the mathematical content involved.

The problem of the converse (see Section 1 above) might also explain the challenge of struggling to find counterexamples. Individuals might provide the characteristics for counterexamples of the converse statement and not the original statement. That means that different challenges might be closely related.

Tabach et al. (2010b) raised the question of whether the mathematical context (e.g., geometry, arithmetic, calculus) had an influence on the difficulties reported when finding counterexamples. In this context, Perkins and Salomon (1989) had already suggested that the context played an important role when producing counterexamples. Zaslavsky and Ron (1998) found out that among geometry and algebra statements, top level high school students made much more reference to counterexamples for geometry statements. For example, more than 75% of the in-service secondary school mathematics teachers in Peled and Zaslavsky's (1997) study managed to produce adequate counterexamples for two false *geometry* universal statements.

Given the close relation between universal and existential statements through negation, it is important to include some of the challenges that might arise when discussing proving and disproving existential statements. In the following section I focus on those issues.

6. Lack of understanding about proving existential statements

In this section I consider two challenges from the existing literature that are linked to the understanding of proving existential statements (ESs): the rejection of one confirming example to prove an ES (Section 6.1) and the rejection of non-minimal proofs for ESs (Section 6.2).

6.1. Rejection or acceptance of one confirming example to prove an ES

Some individuals may not accept that one confirming example is sufficient to prove an existential statement. For example, Tirosh and Vinner (2004) reported that two thirds of the prospective elementary teachers and one half of the prospective middle school teachers who participated in their study claimed that confirming examples could not be regarded as mathematical proofs. Barkai et al.'s (2002) research with elementary school teachers showed that all (27) participants correctly judged an (universally true) existential statement to be true and 50% of them used one or more confirming examples to justify it; however, only half of those considered that those examples were mathematical proofs. In addition, for another case of an existential statement (a not universally true ES), 68% of the teachers identified that it was true; 40% of them provided confirming examples to justify their answer and only 28% regarded their justifications as mathematical proofs.

Possible explanations

An overgeneralization of a proof scheme for USs may explain why some individuals also expect general proofs for the case of ESs and as such they reject confirming examples. Tirosh and Vinner (2004) explained that those teachers seem to have "developed a general view that a mathematical statement is true only if it holds for 'all cases'" (p. 360).

On the other hand, it is possible that those who accept/produce confirming examples to prove an ES had received previous instruction on the sufficient mathematical evidence to prove those cases. They might have learned that in an authoritative way. For example, in their study, Tabach et al. (2010b) asserted that the majority (at least 72%) of 50 in-service secondary school teachers provided a single confirming example to prove the two (true) existential statements they were given. Further, they all correctly judged that one supporting example was sufficient to prove an existential statement. It might be that some of those teachers learned that one confirming example suffices to prove an ES as a rule.

6.2. Rejection or acceptance of non-minimal proofs for ESs

Some teachers do not produce or accept non-minimal proofs when proving existential statements. Tsamir et al. (2009) reported about the case of one secondary teacher who was aware that in order to prove an ES, *one* confirming example was sufficient mathematical evidence. When evaluating non-minimal justifications for true ESs, she rejected them. For instance, for the case of an ES that was universally true, one non-minimal justification had the form of a general symbolic proof; whereas that for the case of an ES that was not universally true, a non-minimal justification had the form of several confirming examples. From the teacher's perspective, those non-minimal justifications involved "overdoing" parts of the justification, which might have revealed the prover's lack of understanding of what suffices to prove ESs.

In contrast, there are cases of individuals who instead of rejecting non-minimal proofs for ESs, produce them. That might also indicate a lack of understanding of proving ESs as the individuals might have overgeneralized a proof scheme for USs to the case of ESs. For example, Tabach et al. (2010b) reported that around 25% of the in-service secondary school teachers in their study provided a symbolic non-minimal proof for two true ESs (one of which was universally true).

Possible explanations

Individuals may reject non-minimal proofs for ESs because they do not satisfy the standard requirements that make an existential statement true. As in Tsamir et al.'s (2009) study, individuals might presume that the "overdoing" goes against the sufficient evidence to prove an ES and may reflect a lack of understanding of proving ESs.

On the other hand, Tabach et al. (2010b) gave three possible reasons for why the teachers in their study produced non-minimal proofs for a true ES that was universally true: first, the teachers "merely reproduced the same general proofs as they had for" the respective USs; second, because of the teachers' curiosity they explored the statements generality and so whether they were true in general or not; third, the teachers could have assumed that as the statement was universally true, then it followed "existentially" true. Although the two last reasons may exhibit some sort of sophisticated understanding of proving and its nature, the first reason may reveal a lack of understanding. It is possible that an individual learned as a rule that proving involves producing general proofs, without making a reflective distinction between existential and universal statements.

Again, proving is closely related to disproving and the case of ESs is not an exception. In the following section I pay close attention some of the struggles that might emerge when individuals are engaged in disproving ESs.

7. Lack of understanding about falsity and disproving of existential statements

In this section I consider two sub-sections that refer to two common challenges reported in the existing literature that reflect a lack of understanding of falsity and disproving ESs: the lack of understanding of what makes an ES false (Section 7.1) and the use of "counterexamples" to disprove ESs (Section 7.2).

7.1. Lack of understanding of what makes an ES false

Individuals may not be aware of or may not understand what it means for an ES to be false. That can be manifested, for example, through their inaccurate judgement of the truth value of ESs. Around half of the prospective elementary teachers in Tirosh and Vinner's (2004) study and 20% of the prospective middle-school teachers "incorrectly argued that the existence theorems that were included [we]re false" (p. 360). Tall (1977) found that the majority of students who had just began their university studies claimed that the (true) statement *"Some rational numbers are real"* was false. However, in a follow-up questionnaire, where S was defined as the set made of nine prime numbers, with two of them (2 and 19) more easily recognizable as prime than the others (257, 601 and bigger ones), the (true) statement *"Some numbers in the set S are prime"* was determined to be true, whereas the (true) statement *"Some numbers in the set S are even"* was regarded as false by most of students.

There is also evidence that even though some individuals produce suitable disproofs for existential statements, they may be unsure about their validity. For example, a few of the in-service elementary school teachers who participated in Barkai et al.'s (2002) study correctly identified that an ES was false and provided a valid justification. Yet, some of those teachers were not sure whether their justifications refuted the statement.

Possible explanations

It is possible that a lack of attention to false existential statements in educational settings explains poor performances when disproving ESs. Tabach et al. (2010a) claimed that "[e]xistential, false statements are rarely addressed in school mathematics" (p. 1086). As shown by Tall (1977), this is not an exception at higher educational levels.

The different way certain words are construed in mathematics in contrast to common-life might influence individuals' assumptions about the falsity of ESs. Tall (1977) pointed out that,

> [a] coloquial interpretation of the word 'some' in many cases seems to be that it means 'two or more of a given set, but without the certain knowledge that it is all of the set.' Shades of meaning vary from person to person. (p. 5)

Indeed, in everyday speech, the statement *"Some A are B"* implies that *"Some A are not B"*; however, in mathematics/logic that inference is not valid (e.g., Newstead & Griggs, 1983; Woodworth & Sells, 1935). Lee and Smith (2009) explained that, "from a linguistic perspective, the uses of these terms often follow the maxims of pragmatics in everyday conversation" (p. 22). Notably, they pointed to the *quantity* maxim (see Grice 1989, and Chapter 2, Section 1.3) as a way to explain the use of "some" in everyday conversations: "It is also informative to the best of the speaker's knowledge, or 'all' would be used instead" (Lee & Smith, 2009, p. 22).

In the context of complex statements (e.g., statements with more than one quantifier) there is also a link between language and dis/proving. Shinno, Miyakawa, Mizoguchi, Hamanaka and Kunimune (2019) evidenced that the way EA statements[5] are interpreted in natural language in Japanese influences their disproving. In their study, 47 undergraduate students were divided into two groups of 24 and 23 students, respectively. The former group received one AE statement[6] and one EA statement in their mathematically symbolic form; whereas the latter group received exactly the same statements, but in their natural-language (verbal) form in English. The students were requested to translate the respective statements to Japanese, determine whether the statements were true or false, and justify their answers. Shinno and his colleagues showed that while the students who received the mathematically symbolic form for the EA statement succeeded in their analysis, most of the students who received the natural-language EA statement, in English, failed to interpret the statement; that is, they failed in translating the statement. Instead, they gave an AE statement and so its truth-value and respective proof were affected given that they relied on an unsuitable transformation. The authors claimed that the way an existential statement is expressed in Japanese, which is different from that in English, might have been the reason for failure in the second cluster of students.

7.2. Use of "counterexamples" to disprove ESs

[5] An EA statement is a statement of the form *"There exists an x such that for every y, $f(x,y)$"*.

[6] An AE statement is a statement of the form *"For every x, there exists a y such that $f(x,y)$"*.

Some individuals consider that "counterexamples" disprove existential statements. In Barkai et al.'s (2002) study most of the elementary school teachers who incorrectly claimed that a (true) existential statement was false, justified their claim by providing what they assumed to be "counterexamples" for it (that is, for a statement of the form *"Some X are Y"*, an *X* that is not *Y*). Similarly, though at school level, Buchbinder and Zaslavsky (2009) reported the case study of a tenth grader. The student was hesitant about the coexistence of both confirming and non-confirming examples for (true) existential statements. At first such a fact led her to conclude that the statements were both true and false.

Possible explanations

Possibly, an overgeneralization of a disproof scheme for USs might lead some students to use "counterexamples" to disprove ESs. Barkai et al. (2002) pointed out that the teachers who resorted to "counterexamples" for ESs in their study "overgeneralized a scheme that holds for refuting 'for all' propositions and used it for refuting existence propositions as well" (p. 63). Similarly, Buchbinder and Zaslavsky (2009) interpreted the pupil's inconsistencies as a result of a generalization she made based from the case of universal statements and her lack of awareness about the distinction between universal and existential statements as she evaluated existential statements as universal statements.

In general, individuals' background or even their perspective of mathematics might explain their performance when proving or disproving mathematical statements. Lin and Tsai (2012) concluded that the future teachers with a mathematical background in their study were more successful and flexible when using valid modes of argumentation to dis/prove universal/existential statements. Only one of the teachers who did not have a mathematical background succeeded in dis/proving the six statements included.

Universal and existential statements are related through negation, and falsity is directly linked to negation. In the following section I pay attention to this relationship, its importance, some methods of negation and the challenges that may be involved in negation.

8. Lack of understanding about negation

There is a close relationship between falsity and negation. Wason (1968) explained it in the following terms:

> The semantic concept of falsity is logically equivalent to the syntactic concept of negation, and it has been shown that both cause difficulty when sentences have to be evaluated or constructed (p. 274)

Because of the link between falsity and negation, understanding negation plays an important role when dis/proving statements (e.g., Barnard, 1995; Dubinsky, Elterman & Gong, 1988; Epp, 1998, 2003; Lin, Lee & Wu, 2003; Sellers, 2020). Epp (2003) pointed out some reasoning principles that she found to be important when disproving universal statements. For example: "one must have some, perhaps, unconscious, awareness that the negation of a universal statement is existential" (p. 887). Epp remarks that formulating negations is important "because to be able to evaluate whether or not a statement is true, one must understand what it would mean for it to be false" (p. 896). In the same line, Dubinsky et al. (1988) stated that negation

> is extremely important in mathematics, for example in making proofs by contradiction or looking for counter-examples. At a conceptual level, when

> complex logical statements are used to define a concept, it may be said that in order to understand what something is, it is essential to understand what it is not. (p. 46)

There is a growing trend to develop research at the undergraduate mathematics level about students' understandings of logical structures of mathematical statements for example in Transition-to-proof or Calculus courses. For example, Sellers (2020) focused on students' quantifications and negations in the context of Calculus statements with multiple quantifiers.

Among the three methods of negation that Dubinsky et al. (1988) discuss, *negation of the meaning* is the method that resorts to this relation. It consists of having "a mental representation of a set of situations that correspond to the statement being true and [can] then take the complementary set of situations which corresponds to its falsity" (p. 46). The other two methods imply in a way the use of rules. *Negation by rules*, the authors explain, is the most mechanical method as it involves retrieving rules from memory. They explain that "[t]he rule for negating a single-level quantification is to replace the universal (alt. existential) quantification by an existential (alt. universal) quantification and to replace the Boolean function by the same function but with its truth reversed" (p. 47). *Negation by recursion* is "transitional towards using an understanding of the statement" (p. 46) and requires identifying the main operation in the statement, negating it either by rules or by meaning, and then proceeding to negate the secondary parts. From the three methods, negation of the meaning is "[t]he most powerful, although most difficult, method of negation" (p. 46).

Challenges when negation is involved

It has been pointed out that an inadequate understanding of negation can be an obstacle to understanding indirect proving (e.g., Antonini, 2001; Lin et al., 2003; Pasztor & Alacaci, 2005). Bardelle (2013) claimed that "[n]egation is a fundamental concept for the construction of meaning in general, and, in particular, for the construction of meaning in mathematical context" (p. 65). She referred to the importance of understanding negation as an important factor when learning other mathematical concepts and understanding the link among them.

Among the many challenges reported in the existing literature, there are several assumptions that individuals use when asked to negate a statement that may reveal their lack of understanding of negation in mathematics. For example, some students assume that negating a statement involves negating only one part of the statement (e.g., Barnard, 1995; Dawkins, 2017; Dubinsky et al., 1988; Lin et al., 2003; Pasztor & Alacaci, 2005; Sellers, 2018). Dubinsky et al. (1988) stated that it seems to be especially the case when statements are complicated.

> In general, students appear to have considerable difficulty with negation. When the statement is complicated, the student will often focus on negating one part or one value of the variable and not consider the entire statement. (p. 47)

Sellers (2018) found out that some undergraduate students negate only one of the variables involved in the statement in order to obtain the negation of *complex mathematical statements*, defined by Sellers as statements that "have two or more quantifiers and/or logical connectives". Similarly, Barnard (1995) identified that the most common error found in his study was the negation of just one part of the statement; however, he also observed that some students negated multiple unrelated components of statements. Barnard (1995) investigated undergraduate students' negation of mathematical and everyday context statements. Two groups of students, 78 first-year and 78 second/third-year university students, were given three sections of seven statements each, which had the same logical structure across sections. The two first sections (one with everyday and the other with mathematical statements) asked the students to choose the negation among a list of four or five options, while the third section provided seven everyday statements for which the students were requested to write the respective negations. Even though the author mentioned that before the participants engaged in solving the tasks they were reminded of the meaning of the word "negation" and received a list of examples, he did not specify what the content of this information was. Similarly, Pasztor and Alacaci (2005) found was that students tend to negate existential statements of the form *"Some X are Y"* as *"Some X are not Y"*, and vice versa (negate *"Some X are not Y"* as *"Some X are Y"*).

Notwithstanding, similar difficulties with negation have been reported as well with the case of simpler (single-) quantified statements. Some students tend to resort to the contrary opposition as if it were the mathematical/logic negation of a single-quantified statement. That means that the negation of universal statements of the form *"All X are Y"* are taken as their contrary statement *"No X are Y"*, and vice versa, the negation of *"No X are Y"* as *"All X are Y"*. Pasztor and Alacaci (2005) studied 32 undergraduate college students' and 15 master and PhD students' negations of single-quantified statements. The authors concluded that the semantic content did not really play a significant role in the students' negations. From their qualitative analysis, the authors determined the students' schemes of negation. For the case of the undergraduate students the highest tendency relied on the use of contraries before instruction, which notably decreased after instruction; whereas the rate of contraries usage for the master and PhD students was remarkably high. The students were given 16 statements, which were divided into four categories: abstract statements (e.g., "All x's are y"); nonsensical statements (e.g., "No borogove is mimsy"); meaningful statements with true negations (e.g., "Some machines are alive"); and meaningful statements with false negations or negations that provide weaker true information than contraries (e.g., "No mammals breathe"). This means that these categories were differentiated in terms of their semantic content. Every category included four statements, which differed in their logical structure; that is, one category included one statement for each of the following forms: *"all…are…"*, *"no… are…"*, *"some… are…"* and *"some… are not…"*. The instrument was applied to the first group of students before and after instruction on negation in order to identify possible changes, whereas the second group only solved the instrument in one opportunity as they all had previously taken a logic class. The students were requested to write the negation of each statement. The "*negation* of a sentence A was defined as a contradictory sentence B that denies the truth of the given sentence A, and in every situation, one of A and B must be true and the other false" (p. 1716).

In addition, some single-quantified statements may be more difficult to negate than others. Lin, Lee and Wu (2003) investigated (202) 17-20-year-old students' negations of statements without quantification as well as "all-", "some-" and "only-one-"[7] statements. They noticed that the students succeeded the most when negating "some-statements" and statements without quantifiers. For those statements, they observed that the context played a role as the real-life statements were about 10% higher in success than the mathematical ones. In contrast, "all-" and "only-one" statements had the lowest rates of

[7] For example, the authors included the case of the statements *"The graph of function f(x) intersected x-axis at only one point"* and *"Engle has only one brother"*.

success, and the authors did not find big differences between the success rate for real-life and mathematical contexts in those cases. The quantitative analysis of Pasztor and Alacaci's (2005) study with undergraduate college and postgraduate students exhibited similar results, where more difficulties when negating universal ("all" and "no") statements than with "some-statements" were found.

Possible explanations

There seems to be an intrinsic difficulty with negative statements in comparison with affirmative ones that may explain the challenges that arise when negation is involved. In this respect Horn (2001) has pointed out that

> negative statements are harder to verify than their affirmative counterparts, that the difficulty posed by negation correlates directly with the implausibility in the context of the corresponding affirmative supposition, and that overt negation presents more problems to the language processor than does inherent or implicit negation. (p. xvi)

Some of these difficulties with negation may be explained by the students' use of memorized rules for negation (Dubinsky et al., 1988). For instance, learning to formally negate statements by rules may explain why some students apply rules mechanically without understanding them.

It is not uncommon to expect that once students learn rules to negate statements, they resort to those rules without attempting to make sense of them. In Pasztor and Alacaci's (2005) study, some students used a *mechanical response* scheme, which is similar to what Dubinsky et al. (1988) called *negation by rules.* The undergraduate students increased their use of the mechanical responses scheme after intervention which, according to the authors, suggested the students' application of syntactic rules. For example, even though some students correctly negated statements of the form *"Some X are Y"* as *"All X are not Y"*, they did not directly give *"No X is Y"*. Similarly, other students correctly negated statements of the form *"Some X are not Y"* as *"No X is not Y"*, instead of directly providing *"All X are Y"* (equivalent to *"No X is not Y"*). Pasztor and Alacaci concluded that the high increase of this sort of negations "suggests that instruction of formal negation may actually foster a mechanical approach to negation" (p. 1719), although they did not specify whether the students were given rules to negate statements during instruction.

Another common assumption that many students use, Durand-Guerrier et al. (2012) claim, is that the mathematical/logic negation of an "if-then" statement is also an "if-then" statement. However, in mathematics the negation of the statement *"if p then q"* is not a conditional statement, but the conjunction *"p and not q"*.

The everyday-life context might also explain many of the challenges with negation. Notably, some may stem from the way students interpret quantifiers and logical connectives, which are sometimes different from mathematical conventions (e.g., Dawkins & Cook, 2017; Dawkins & Roh, 2016; Epp, 1999, 2003; Selden & Selden, 1995). Indeed, the approaches that individuals without any previous instruction about negation spontaneously use to negate quantified statements usually differs from the way it is expected in mathematics. Most likely their only source of reference is their everyday language framework and its conventions for negation. For example, Epp (2003) made that point for the case of conditional statements. She emphasized that,

> [o]rdinary language contains many different varieties of if-then statements besides the mathematical kind – ones referring to causal relationships, temporal relationships, counterfactual situations, and so forth. There are conventions for

> negating if-then statements in these other situations, but they are different from the conventions of mathematical logic. (p. 889)

Epp (2003) provided some examples to illustrate her observations, where she precisely shows that "if-then" statements are negated as "if-then" statements in informal language.

> Imagine that a friend states "If I were Ann, I wouldn't do what she did" and we disagree. We might well say, "No, if you were Ann, you would do exactly what she did" … Or to counter the claim that "If carbon emissions continue to occur at the present rate, the earth's temperature will increase by 10 degrees", we might say "No, even if carbon emissions continue to occur at the present rate, there does not necessarily have to be a 10-degree increase in the earth's temperature". (p. 889)

Ordinary speech may have a big impact in the way individuals negate quantified statements in mathematics, especially if they do not become aware of the differences between different contexts. Epp (2003) stated that "[i]n ordinary English, one can negate a universal or existential statement in several different ways, one of which is simply to insert the word 'not'" (p. 890). As an illustration, Epp provided the statement *"All grass is green"*, and some possible negations in everyday speech: *"Not all grass is green"*, *"All grass is not green"* and *"Some grass is not green"*. This is similar to what happens in colloquial speech in Spanish, even though this is not the way negation works in mathematics[8].

Antonini (2001) hypothesized that one major obstacle to learn mathematical negation was the difference between negation in mathematics and negation in natural language. He claimed that in contrast to mathematics, where it is always possible to express negation in an affirmative form, in natural language this is not the case. To illustrate his point, he provided the example "f is a non-increasing function", which in mathematics can be affirmatively expressed as "there exist x, y such that $x < y$ and $f(x) \geq f(y)$". Another obstacle he referred to is the individuals' tendency to classify objects based on their differences more than on analogies. Antonini exemplified this with the case of a parallelepiped and a hexagonal prism, which can be seen based on their differences and separated into two classes, or the class of parallelepipeds can be seen as included in the general class of prisms in its mathematical classification.

An everyday-context interpretation of the quantifier "some" may explain the way individuals determine the truth value of implicit negations of "all-statements". Bardelle (2013) found out that many Italian science university students made decisions on the truth and falsity for an implicit negation based on its colloquial interpretation. The four tasks Bardelle provided involved the statement *"Not all circles are black"*. The students were given three diagrams (per task) that contained black and/or white circles from which to choose one or more options that made the statement true or false. The students exhibited their use of an interpretation of the statement that is consistent with its everyday-speech interpretation, that is, they chose diagrams where *"Some circles are black"* and *"Some circles are not black"* for the truth cases. Some students decided their falsity cases by relying on the truth cases; that is, they first decided what diagrams made the statement true and then chose the remaining options as the diagrams that made the statement false. Almost none of the students chose the diagrams where all circles were white, which was consistent with their approach based on their colloquial interpretation of the implicit

[8] For an account of negation of SQ-statements in mathematics, see Chapter 2, Section 2.1.

negation. Bardelle underscored that the tasks that analyzed truth cases were more difficult than those for falsity.

Likewise, the fact that the contrary statement of a universal affirmative statement is chosen as its negation might have to do with the way negation is used in ordinary conversation where, as Abrusci, Pasquali and Retoré (2016) pointed out, sentences such as *"Not every picture tells a story"* may be understood as *"No picture tells a story"*.

Durand-Guerrier (2020) discussed the way negation is construed in different contexts in French and how this may affect students whose native language is not French, but learn mathematics in French. She claimed that sentences of the form *"For all A, A is not B"* in French are ambiguous. According to linguistic norms, they mean *"some A are B, and some are not"*, while in everyday life it might mean *"no A is B"*. She illustrated the confusions that may arise with these different interpretations with a daily-life example: "In a very cold winter in Lyon (France), the Public Transportation Company had widespread the following message: 'Aujourd'hui, tous les bus ne circulent pas' (Today, all buses do not circulate). A number of people called to ask which buses circulated and three hours later, the original message was changed to 'Aujourd'hui, aucun bus ne circule' (Today, no bus is circulating)" (p. 34). Durand-Guerrier explained that a non-standard interpretation is more common in oral communication, which includes radio and broadcast. Furthermore, those ambiguities intrinsic to French language seem to be the source of difficulties individuals from a non-French speaking country face when learning mathematics in French. Durand-Guerrier referred to Kilani's study. Kilani pointed out the differences in the grammatical structures between Arabic, French and predicate calculus. Kilani showed that negation is the same in Arabic grammar and in predicate calculus, but it was different in French. He pointed out that this was a problem in Tunisia, where mathematics is taught in Arabic from grade 1 to 9, and in French later than that, given that language and mathematics teachers did not seem to take these differences into consideration. Durand-Guerrier suggested using the logical analysis of statements as a tool to overcome grammatical misunderstandings in mathematics education.

The way some words are used in an everyday context might also influence the way they are used in mathematics. Hempel and Buchbinder (2022) investigated the way prospective secondary teachers used the word "opposite" from their everyday language to describe logical operations in the context of indirect proof, notably, the contraposition equivalence, negation in proof by contradiction and converse. For example, "opposite" was used to justify why a statement and its contrapositive are equivalent (because the contrapositive is its opposite), to substitute the word "negation"; and to create the converse statement. The authors claimed that it is important to identify (correct and problematic) patterns in the pre-service teachers' use of language related to indirect proof and to not ignore them.

The truth value of statements may also have an impact in negation (e.g., Barnard, 1995; Pasztor & Alacaci, 2005). Pasztor and Alacaci (2005) found that the truth value of the statements played a crucial role for the students' negations of "some-statements". The authors called this scheme *truth value effect*. For instance, one student correctly negated *"Some living things don't grow"* as *"All living things grow"*, while she incorrectly negated all her other "some… are/do not…" statements, which were abstract, nonsensical or had a false negation (see above). In the example, the fact that the given statement (*"Some living things don't grow"*) was false facilitated finding a true negation. Another student correctly negated *"Some machines are alive"* as *"No machines are alive"* even though all her other negations of "some… are…" were wrong and it was the only time the student used the quantifier "no". In contrast, some students had correctly negated all statements with exception of the (true) statement *"Some people don't live under water"*, which they (incorrectly) negated as the (true) statement *"No people live under water"* instead of *"All people live under water"*. In Barnard's (1995) study the truth value of everyday-context statements had an influence in the higher number of successful negations for the group of second/third-year students. For instance, a student found it difficult to write down the negation for everyday statements that s/he knew that were true.

Furthermore, the form in which some statements are expressed may also influence their negation. While everyday statements are presented as worded statements, mathematical statements involved the use of symbolic notations, which some students find more difficult to deal with (e.g., Barnard, 1995).

All the studies I included in this first section are important as a reference for knowing where to act in terms of the design of an intervention that has the purpose to improve teachers' poor skills with dis/proving. Lee and Smith (2009) had already suggested that research needed to be developed to "investigate the relation between students' interpretations and proof constructions, and also to see how instructional practices can attend to the cognitive challenge faced by students who are new to 'the rules of the proving game'" (p. 25). In that regard, several studies I mentioned before show that the issue of language has an influence in developing proof-related skills.

In the next section I include some directions that were given by other researchers in terms of the knowledge that teachers might need to engage their students in proof-related activities. Within it, I focus on the interventions that have been developed to try to remediate some of the aforementioned proof-related challenges and that are obstacles for proof-related teaching.

II. Knowledge for proof-related teaching

In order to promote reasoning and proving in mathematics classrooms, it is fundamental that teachers understand different aspects of proof. Some researchers have given some hints on the knowledge that might be important for teachers to have and use when engaging their students in proof-related activities (see Epp, 2020, for many examples).

Among those hints, awareness of logical principles of reasoning has been pointed out as crucial by several researchers. Durand-Guerrier et al. (2012) explained that,

> including instruction in logical principles as one part of mathematics education provides a balance between two extreme positions: that checking the validity of a proof requires that it be completely formalised; and that success in proving requires no explicit knowledge of logical principles. (p. 375)

Indeed, Epp (2003) claimed that "[d]etermining truth and falsity of mathematical statements is so complex that, even when they are motivated, students often fail to 'get it' if they do not have some knowledge and experience with basic logical tools" (p. 895). In particular, she highlighted the importance of understanding quantifiers in mathematics and the differences among them.

> In mathematics the distinction between "all" and "some" is crucially important. Whether a statement begins "for all" or "there exists" completely determines how to tell whether or not it is true and what we can deduce from it. (p. 889)

In agreement with Epp, Durand-Guerrier et al. (2012) suggested that

> [t]eaching should aim to make explicit the assumptions about the uses of logic, in terms of both reasoning and language, that students have not yet learned… However, at any grade level, teachers cannot effectively guide their students' reasoning activities, if they themselves are not explicitly aware of the basic principles of logical reasoning. (p. 377)

Understanding the nature of proof is hinted differently depending on the aspect attended. For example, Jahnke and Wambach (2013) emphasized the importance of focusing on inferences when it comes to proving a mathematical statement. They claimed that it is important to understand that,

> a mathematical proof does not establish *facts*, but *implications*. A proof proves 'if-then-statements'. That means a mathematical theorem is not about a fact *B*, but about an implication 'If *A* then *B*'. Thus, the *absolute certainty* of mathematics does not reside in the fact, but in the *inferences*… any mathematical theorem is to be considered as an 'if-then statement' because its truth depends on other statements. (p. 471)

It has been recommended that teachers should gain experience constructing and evaluating arguments themselves before they engage their students in similar activities. Knuth (2002b) claimed that,

> having teachers construct and present proofs of school mathematics tasks – tasks from various content areas and levels – provides a forum for discussing expectations of proof (e.g., what counts as proof) for students at differing levels of mathematical ability and in different mathematics courses. (p. 84)

In addition, Tsamir et al. (2009) suggested that,

> [t]eachers need to be able to evaluate students' suggested proofs to various statements. That is, teachers need to identify students' correct and incorrect justifications when determining the validity of true and false statements. (p. 60)

As a skill that supports the construction and validation of arguments, Selden and Selden (1995) discussed "unpacking the logical structure of an informal statement". In the same vein, Epp (2009b) suggested translating exercises as an important part of learning to produce and judge the correctness of arguments, which students may find a difficult task. Epp's teaching included "translating back and forth from formal mathematical statements to their many different informal versions". In particular Epp considered using this approach for the case of definitions, because as she claimed, "it is often the case that truth or falsity of a mathematical statement is more apparent if one uses one phrasing of a definition rather than another" (p. 316).

Teachers should also be aware of the different valid modes of argumentation and valid modes of expression (e.g., verbal, generic, visual arguments) in which proof can be presented (e.g., Dreyfus, 2000). Not only that, but it is also crucial that teachers become aware of the differences between common and mathematical language and the way some terms or expressions are interpreted accordingly (see previous sections).

Interventions with a proof-related focus

The design and implementation of interventions that tackle some of the recurrent proof-related challenges is a way to contribute to the field by remediating a possible lack of pre-/in- teachers' background related knowledge to engage their students in proof-related activities. In that respect, several classroom-based interventions have been designed and implemented in the past years (e.g., Buchbinder, Zodik, Ron & Cook, 2017; Harel, 2001; Jahnke & Wambach, 2013; Mariotti, 2000, 2013; G. J. Stylianides & A. J. Stylianides, 2009, 2014).

Some interventions have been designed to tackle specific, common and persistent proof-related challenges. For example, G. J. Stylianides and A. J. Stylianides (2009) focused on the common assumption students make that empirical arguments do qualify as proofs (see "possible solutions" in Section 2.1 above).

Some other interventions aimed more broadly to identify the types of knowledge/understandings that are relevant for teachers to develop to successfully engage their students in proof-related activities. For instance, A. J. Stylianides and Ball (2008) pointed out that most of the research into pre-/in- teachers' knowledge about proof has explored their understandings of logico-linguistic structure of proof, which encompasses understanding the role that language plays in proving, understanding and distinguishing empirical and deductive arguments, etc. Prior research has exhibited several difficulties teachers have in this area (see Section I in this chapter for some common challenges teachers might face).

> This body of research has shown that teachers of all levels tend to have weak knowledge about the logico-linguistic structure of proof …, thereby suggesting the inadequate preparation of many teachers to effectively cultivate proving in their classrooms. (A. J. Stylianides & Ball, 2008, p. 329)

Stylianides and Ball investigated the type of knowledge about proof that is likely to be important for teachers to engage their students in proving. They observed that within the knowledge about the logico-linguistic structure of proof, little (if any) has been explored in terms of research with focus on teachers' understandings of systematic enumeration for the case of a finite number of cases involved when proving. Further, Stylianides and Ball suggested that the knowledge about the logico-linguistic aspects of proof should be complemented with what they called "knowledge of situations for proving", for what they included *knowledge of different kinds of proving tasks* and *knowledge of the relationship between proving tasks and proving activity*.

A. J. Stylianides (2011) contributed to A. J. Stylianides and Ball's (2008) knowledge package for teaching proof by identifying two important elements that, as he claimed, would be required for other teachers to implement G. J. Stylianides and A. J. Stylianides' (2009) intervention in their classrooms; namely, (1) understanding the organization of the activities included in the intervention design to support students' progression along the intended learning trajectory, and (2) understanding the rationale for considering the conceptual awareness pillars included in the intervention so that the students are more aware of their existing conceptions and so they experience the expected cognitive conflicts.

G. J. Stylianides, A. J. Stylianides and Shilling-Traina (2013) investigated the challenges that prospective elementary school teachers faced when teaching reasoning-and-proving (RP) in their mentor classrooms. In order to possibly avoid the presence of previously reported challenges (i.e., weak mathematical knowledge about RP and counterproductive beliefs about its teaching), the authors chose participants who had participated and stood out in a prior study that aimed at developing their knowledge about RP (see G. J. Stylianides & A. J. Stylianides, 2009). Some of the challenges pointed out by those prospective teachers were related to mentor teachers' classrooms (e.g., challenges related to classroom norms and students' habits of mind), their lesson planning and implementation (e.g., those related to high-level tasks and sense of effectiveness); and their own knowledge (e.g., their knowledge of mathematics, students and the curriculum). Moreover, the three most common challenges as alluded by the future teachers, were their implementation of high-level tasks, the students' habits of mind in mentor teachers' classrooms and their knowledge of students.

Despite all those efforts, G. J. Stylianides, A. J. Stylianides and Weber (2017) considered that "[e]xisting research on classroom-based interventions has barely scratched the surface of a few of the open problems in the teaching and learning of proof". Among some of the relevant examples of issues that the authors consider that still need to be addressed in this context are: "[r]esearch at the school level on the teaching of the axiomatic structures of mathematics", which the authors observed that most of it has been developed in secondary school geometry contexts, and they added that "[t]here is a great need for interventions to introduce students of different ages to an appropriate set of criteria for judging whether an argument meets the standard of proof" (p. 257).

More recent research has explored various aspects of prospective teachers' Mathematical Knowledge for Teaching Proof (MKT-P). For example, Buchbinder and McCrone (2020) divided the MKT-P into two types of knowledge: Subject Matter Knowledge related to Proof (SMK-P) and Pedagogical Content Knowledge related to Proof (PCK-P). On one hand, for the SMK-P they considered the Knowledge of the Logical Aspects of Proof (KLAP), within they included knowledge of various types of proof (e.g., proof by contradiction, direct proof, proof by cases, disproof by counterexample), in/valid modes of reasoning, logical relations (e.g., negation, conjunction, biconditional), various definitions, multiple methods of proof (i.e., various proofs for one theorem). In my view, Buchbinder and McCrone's KLAP resembles to what Stylianides and Ball (2008) called the knowledge about the logico-linguistic aspects of proof. On the other hand, the PCK-P was constituted by Knowledge of Content and Students (KCS-P) and Knowledge of Content and Teaching (KCT-P), which may be seen as including what Stylianides and Ball (2008) called knowledge of situations for proving. Buchbinder and McCrone put together a course for senior future secondary teachers, which was organized in terms of modules that were aligned with four proof themes: (1) quantified statements; (2) conditional statements; (3) direct proof; and (4) indirect reasoning. As part of each module, the future teachers were expected to enhance their KLAP, examine students' proof conceptions related to the specific theme, plan a lesson with that theme as a frame, teach the lesson to a small group of secondary students, reflect on their work and, write down a report. Buchbinder and McCrone reported on the planning, implementation and reflection of preservice teacher's lessons on the second theme, conditional statements. They found out that in general the most common pedagogical feature that the prospective teachers used for their lesson planning was analyzing others' work or arguments. The higher number of lesson plans with a high focus on the respective proof theme were those for conditional statements and direct proof; whereas quantification and indirect reasoning had the higher number of lesson plans with a low focus. In terms of their implementations, the future teachers' attempts to make explicit the mathematical/logic ideas about conditional statements was affected by their use of appropriate mathematical language. The preservice teachers found it challenging to find the right balance between using accurate and consistent language while adjusting it to the students' level. As a whole, Buchbinder and McCrone (2020) concluded that there were two types of challenges that the prospective teachers faced:

> (a) the PSTs' [preservice secondary teachers'] own doubts about utility and feasibility of teaching proof themes at the secondary level, and (b) challenges specific to planning and enacting lessons on the four proof themes. (p. 14)

Some prospective teachers pointed out that they were concerned about teaching the four proof-related themes in a subject different from Geometry, or teaching those to low achieving students, although many of them manifested an optimistic view on the latter matter. The preservice teachers were not sure about what learning proof-related goals their students could achieve in only one lesson. Some future teachers considered some proof themes more challenging to integrate than others.

Buchbinder and McCrone (2022) built on the concept of proof-based teaching (see Chapter 1) to suggest three guiding principles for teaching mathematics via reasoning and proving. The first principle they suggested was to embed the teaching of reasoning and proving within the content that is part of the mathematics curriculum. The second guiding principle underscored the use of deductive reasoning to construct and validate mathematical knowledge. The third principle pointed to the use of language, notation and representation that is appropriate to the students. Buchbinder and McCrone designed a course for future secondary teachers, who designed four proof-oriented lesson plans based on those principles and implemented the lessons in schools.

G. J. Stylianides and A. J. Stylianides (2009) used the cognitive conflict approach for the design of their intervention that aimed at preservice teachers' realization of the limitation of empirical arguments as valid methods to prove USs. In the following section I explore some of the factors that should be considered when the approach is used, some situations that can support triggering cognitive conflicts, and what other studies related to proof have adopted the cognitive conflict approach.

III. The Cognitive Conflict Approach and Proof

In Chapter 3 I outline some important theoretical elements of the cognitive conflict approach (see Sections I.1.1 and I.1.2). As I mentioned before, here I focus on research related to what to consider when using the approach as they might influence the individuals' responses to the approach (Section 1), what conditions can support the approach (Section 2) and its use in research related to proof (Section 3).

1. Factors to consider when using the cognitive conflict approach

The cognitive conflict approach has been considered as a strategy that has shown positive effects in learning (e.g., Limón, 2001; Okazaki & Koyama, 2005; Ngicho et al., 2020; Steffe, 1990; Swan, 2001; Tirosh & Graeber, 1990; Zaslavsky, 2005). Nonetheless, caution is recommended when using cognitive conflicts or when evoking uncertainty.

Zaslavsky (2005) pointed out that frustration, application of erroneous procedures to resolve the uncertainty, and a resistance to change some initial rules are some of the behaviors that may be triggered when uncertainty is evoked. Gal (2019) observed that ineffective use of the cognitive conflict strategy might be explained in terms of the teachers' lack of awareness about their students' necessary knowledge of concepts and principles, as well as their ability to bring those into practice. In her paper, Limón (2001) reported on different positive effects of using the cognitive conflict approach; however, she also discussed other factors that should be considered when meaningful conceptual changes are aimed, since "[s]ometimes, partial changes are achieved, but in some cases they disappear in a short period of time after the instructional intervention" (p. 364). The factors she considered are: motivational factors, students' prior knowledge, students' epistemological beliefs, students' values and attitudes, students' learning strategies and cognitive engagement, social factors, and students' reasoning abilities.

In terms of factors that may influence how individuals respond to anomalous data, Chinn and Brewer (1993) discussed four: individuals' prior knowledge; a possible alternative theory; the anomalous data; the processing strategies that guide the evaluation of the anomalous data. For an *individual's prior knowledge*, the authors considered that "the more entrenched a belief, the harder it should be to persuade an individual to change the belief' (p. 15). In that respect, the authors point out that special attention should be given to ontological and epistemological beliefs given that they are "relatively immune to change" (p. 17). Chin and Brewer also considered that individuals' background knowledge can influence the way they respond to anomalous data. "On one hand, background knowledge can lead an individual to reject or reinterpret anomalous data. On the other hand, background knowledge can lead the individual to accept the anomalous data and to make either peripheral or core changes to the current theory" (p. 18). The authors' position was that "subjects will reject or reinterpret anomalous data if they can access relevant background knowledge" (p. 20); however, there are also cases where individuals lack of background knowledge, which Chin and Brewer claim that can facilitate theory change. There might be cases where the individuals have too little background knowledge, the authors acknowledge, that they do not understand the anomalous data, which might lead them to ignore or exclude the data. Another factor involves whether there is an *available plausible alternative theory* that count on a mechanism to explain the new data and how good the new theory is. There are some characteristics of *the anomalous data* that may influence a response to it. For example, it is important to consider whether the data is credible; otherwise, it will be immediately rejected. Similarly, if the data is not ambiguous, it can increase the chances for a theory change. On the other hand, if the data is ambiguous, it may favor reinterpretations to keep their initial theory. Additionally, the *strategies for processing the anomalous data* can also affect the way an individual respond to it. Theory change is more likely when an individual processes anomalous data deeply, the authors claim. That may be encouraged, for instance, by "choosing an issue that is personally involving to the reasoner (…) or tell reasoners that they will have to justify their reasoning to other people" (p. 29).

2. Situations where cognitive conflicts may arise

Becoming aware of the situations that may create a cognitive conflict can support the design of interventions that use the approach to foster new understandings. Some studies report on situations that may trigger cognitive conflict (e.g., Meissner, 1986; Zaslavsky, 2005).

Meissner (1986) included four situations where conflicts may occur: U-shaped behavioral growth, gaps, defective frames, and wrong frames. Meissner explains that the first seems to be age-dependent and is observed in younger children. *U-shaped behavioral growth* refers to a situation where an initial spontaneous concept that is based on intuitive thinking interferes with a posterior scientific or cultural concept that relies on analytical thinking and is developed through schooling. The children's initial global view turns into "chaos" and children cannot solve problems they could solve before. Schooling constructs new scientific concepts that serve as the basis for some of their initial abilities which might reappear later, whereas other abilities disappear. *Gaps* refers to situations where there are two views of a problem: an informal and a formal view. For example, when a child correctly reduces the fraction $\frac{20}{35}$ to $\frac{4}{7}$, but when asked to choose to get $\frac{20}{35}$ or $\frac{4}{7}$ of a cake, he prefers $\frac{20}{35}$ as he thinks there are more pieces in such a case. Children do not see the conflict between the different results because for them those belong to different context problems. *Defective frames* refers to situations where a rule that applies to a certain context is pushed further and applied in a different frame, where the rule is wrong and make the frame defective. For example, a defective frame for division may include the rule "zero means nothing and adding zero doesn't change it" and an application of such a rule to the division 3606 divided by 2 may lead to conclude that result would be the same as that from dividing 366 by 2. *Wrong frames* refers to situations where "there is a mismatch of understanding because of thinking in different frames" (p. 13). The frame that is activated to solve a problem is not the one as expected by the teacher, which leads to a different answer, usually a wrong answer from the teacher's perspective.

Zaslavsky (2005) focused on the broader term "uncertainty" to encompass the interrelated constructs "conflict, doubt and perplexity" (p. 300). She reflected on the design and implementation of mathematical tasks that evoked uncertainty. Zaslavsky described three types of uncertainty entailed in certain mathematical tasks: competing claims, unknown path or questionable conclusion, and non-readily verifiable outcomes. *Competing claims* include "outcomes, definitions, beliefs, a priori expectations, assumptions, and assertions" that are in conflict with another claim about the same object. She illustrates when this type of uncertainty may be evoked; namely, a discussion on whether to define $(-8)^{1/3}$ as $(-8)^{1/3} = \sqrt[3]{(-8)} = -2$ or $(-8)^{1/3} = (-8)^{2/6} = \sqrt[6]{(-8)^2} = \sqrt[6]{64} = 2$. Zaslavsky explains that "uncertainty raised by competing claims need not be connected to correctness" (p. 301). *Unknown path or questionable conclusion* is a type of uncertainty evoked by exploration tasks and open-ended problems, where students may be uncertain of what exactly to expect. These tasks are common in technologically enhanced learning environments, and existential tasks as "they present the need to check if certain cases are at all possible" (p. 302). *Non-readily verifiable outcomes* "has to do with the lack of confidence one may have regarding the correctness or validity of an outcome (e.g., solution to a problem) which requires verification" (p. 304). Zaslavsky claims that this type of uncertainty is commonly presented in probability and combinatorics tasks, where students tend to be less certain about their outcomes. She continues, "if one does not know how to verify a solution to a problem, he or she may feel a sense of uncertainty, even if there are no competing solutions at question" (p. 305). Zaslavsky noticed the dynamic between these types of uncertainty and the iterations involved in task design.

3. Use of the cognitive conflict approach in proof-related research

This approach has been deployed in studies with several mathematical topics as a focus. For example: decimals (Okazaki & Koyama, 2005; Tirosh & Graeber, 1990), infinity (Tsamir & Tirosh, 1999), and prime numbers and divisibility (Zazkis & Chernoff, 2006, 2008).

In the field of proof and proving, provoking uncertainty and cognitive conflict has been regarded as one way to create a need to prove (Zaslavsky et al., 2012). In particular, Zaslavsky and her colleagues discussed three research studies with such a focus (Hadas et al., 2000; G. J. Stylianides & A. J. Stylianides, 2009; Zaslavsky, 2005).

Hadas et al.'s (2000) study took place in the context of a Dynamic Geometry Environment (DGE). Hadas and her colleagues used two activities that aimed at changing students' initial conjectures when contrasted with their findings supported by the DGE. Particularly, the second activity created uncertainty on whether certain geometrical constructions were possible. The two activities involved the use of cognitive conflicts and were designed with the purpose that students have the need to provide deductive arguments. The authors showed that the approach they used revealed an increase in the percentage of deductive explanations that students provided from one activity to the other. The second activity had a higher percentage of deductive explanations, which may be due to the need to explain the existence of a construction, by showing it, or the impossibility, by deductively explaining why was the case.

G. J. Stylianides and A. J. Stylianides (2009) used the cognitive conflict approach in their design of an intervention that aimed at pre-service teachers' realization that empirical arguments are insecure methods of validation and create a need to learn about secure validation methods. G. J. Stylianides and A. J. Stylianides (2014) pointed out two conditions that were crucial to their design of such an intervention. First, "[a] counterexample is more likely to create a cognitive conflict for students in the area of proof if it is in accord with students' immediately accessible conceptions about proof" (p. 395); and second, "[a] counterexample is more likely to create a cognitive conflict for students in the area of proof if students become more aware of their conceptions about key issues related to proof that the counterexample aims to challenge" (p. 396).

Zaslavsky (2005) made a summary of theories that are the grounds for creating learning situations that entail uncertainty. She highlighted the importance of including tasks that provoke uncertainty given that they foster a wide range of mathematical activities, from which she particularly called attention to two proofs by contradiction that emerged spontaneously during an in-service workshop. She found out that those indirect proofs turned out to be more convincing than a direct method that involved auxiliary constructions in the involved geometry task. Zaslavsky also pointed out that uncertainty tasks created a need for explanatory proofs.

Additionally, Zaslavsky (2008) claimed that teachers can see the value of uncertainty and cognitive conflict to promote a need to prove if they engage themselves in solving tasks that trigger uncertainty and cognitive conflict.

Buchbinder and Zaslavsky (2011) used tasks that they called *"Is this a coincidence?"* with top-level high school students and experienced secondary teachers. The task was designed by the authors in such a way that they triggered uncertainty, which they hypothesized that could foster a need to prove. The authors organized the participants' responses to the tasks by their degree of confidence (strong uncertainty, moderate uncertainty, strong confidence) that was evoked when engaged in solving the tasks. Buchbinder and Zaslavsky claimed that this type of tasks "was found to have a potential for creating a need to prove or convince … This need emerged spontaneously as a way to resolve an uncertainty, evoked by the task" (p. 278). Nonetheless, I believe that the need to prove did not really stem from the uncertainty evoked by the task, as it is claimed by the authors. For example, the authors claimed that Gila, one of the two students in case 3, "successfully refuted this false statement ["A quadrilateral with diagonals that are of equal length and perpendicular to each other is a kite"] by constructing a counterexample" and "in solving other geometry problems … had often mentioned that in geometry it is very important not to rely on visual information" (p. 275). As I see it, this reveals Gila's preconceptions about proving. Notably, it exhibits her awareness that a counterexample disproves a US and that visual evidence is inconclusive when proving is involved. Hence, her need to prove may have been evoked by sources different from the uncertainty the task was aimed to provoke. Gila seems to have used her already-held proof-related assumptions to realize that in order to prove the statement more than what was suggested in the task was expected. Likewise, the case of Natalie, one of the teachers in the paper (see case 6), exhibited a similar related background. In my view, all this is enough evidence to presume that in Buchbinder and Zaslavsky's study the participants already counted on some previous relevant knowledge related to proof and, specifically, with what makes an argument a proof. Most likely their main source for their need to prove had stems from those previous experiences, where they might have learned to prove and about certain principles on what suffices to prove (e.g., they should not only rely on the visual information provided through the figures they are usually shown, as it is only one of possibly infinite cases involved in a mathematical statement).

Buchbinder, Zodic, Ron and Cook (2017) focus on some aspects of the design of a technology-based task ("What can you infer from this example?") that aimed at enhancing pre-service teachers' content and pedagogical knowledge about the status of examples in proving. In such a design the authors considered that the several examples they included would evoke uncertainty and cognitive conflict on two issues: the truth-value of a statement and the status of those examples in dis/proving the statement. They claimed that the chosen examples as well as the order they followed "would eventually support resolving the uncertainty, and serve for PST's [pre-service teachers'] as a *pivotal* (set of) *examples*" (p. 226).

The cognitive conflict approach is a broad teaching strategy that has not been exclusively used in proof-related teaching. Yet, there is a growing interest in using it for the latter purpose. In this section I included some of the factors that are important to have in mind when this approach is used and some of the situations where cognitive conflicts can occur.

Chapter 3: Research Basis

"... it was not till within the last few years that it has been realized how fundamental any and some are to the very nature of mathematics."

A. N. Whitehead (as cited in Epp, 2020, p. 108)

This chapter is divided into two main sections that describe the research basis I used to develop my research. The first section (Section I) includes the theoretical framework and some important constructs from current research that I adopted to guide my work. The second section (Section II) outlines the mathematical framework, which includes the mathematical assumptions I used for the design of the interventions.

I. Theoretical Framework

One of the aims of my research was promoting the teachers' understanding of proof-related issues which, according to the existing literature (e.g., see Chapter 2), probably differed from the teachers' initial assumptions. Considering this implied that I needed to search for approaches that foster change of, or disequilibria, in the system of the teachers' initial assumptions. Section 1 has a focus on constructivism and cognitive conflict as a way to address this aspect. In addition, my interest in the development of mathematical proof-related assumptions obliges me to include a section where I clarify the conceptualization I use for proof (Section 2). As many differences in the teachers' initial assumptions and mathematical assumptions can be explained in terms of the use of language and their flexibility to accept examples as sufficient evidence to prove (for details, see Chapter 2), Section 3 focuses on these topics.

1. Constructivism and Cognitive Conflicts

Constructivism is the main framework of my research. Piaget was one of the first people to discuss constructivism. According to Sjøberg (2010), constructivism has the following core ideas:

- Knowledge is actively constructed by the learner, not passively received from the outside. Learning is something done by the learner, not something that is imposed on him.
- Learners come to the learning situation (in science, etc.) with existing ideas about many phenomena. Some of these ideas are *ad hoc* and unstable; others are more deeply rooted and well developed.
- Learners have their own individual ideas about the world, but there are also many similarities and common patterns in their ideas. Some of these ideas are socially and culturally accepted and shared and are often part of the language, supported by metaphors, etc. They also often function well as tools to understand many phenomena.
- These ideas are often at odds with accepted scientific ideas and some of them may be persistent and hard to change.
- Knowledge is represented in the brain as conceptual structures and it is possible to model and describe these in some detail.
- Teachers have to take the learner's existing ideas seriously if they want to change or challenge these.

- Although knowledge in one sense is personal and individual, the learners construct their knowledge through their interaction with the physical world, collaboratively in social settings and in a cultural and linguistic environment. (p. 486)

Piaget focused on finding out what the nature of knowledge was and how it developed. In that sense, he refuted the empiricist and nativist theories. Piaget did not accept the empiricist and behaviorist idea that knowledge derived directly from sense experiences. Likewise, as Sjøberg (2010) pointed out, Piaget rejected the rationalist or preformist stance that "knowledge is innate and develops more or less biologically as we grow and mature" (p. 486). Piaget spent most of his lifetime showing the insufficiency of those views about the nature of knowledge.

Within constructivism, some important notions play an important role. In fact, I use constructivism to frame the idea of *cognitive conflict*, but also to introduce some related concepts like *equilibration*, *disequilibrium*, *assimilation*, (functional and metamorphic) *accommodation*. My main focus in this section is the cognitive conflict approach.

A cognitive conflict is a situation where an individual faces a mismatch, contradiction or inconsistency between her/his existing assumptions (beliefs, conceptions, knowledge, ideas, etc.) and new ones. Zazkis and Chernoff (2006) add that, a "*[c]ognitive conflict* is an analogue of disequilibration, referring to a pedagogical setting and a learner's cognitive development" (p. 465).

Piaget's equilibration theory is the main theoretical basis for the cognitive conflict approach (Okazaki & Koyama, 2005; Zaslavsky, 2005). The theory of equilibration is framed in *constructivism*. Dubinsky et al. (1988) explain that, according to constructivism,

> what the individual constructs are *schemas*. A schema is a more or less coherent collection of mental objects and mental processes for transforming objects. When faced with a new situation or, what we may refer to in mathematics as a *problem*, an individual is said to be *disequilibrated* and may attempt to *reequilibrate* by *solving the problem*. The process of equilibration results in the construction or reconstruction of schemas. (p. 44)

Piaget (1985, as cited in Okazaki & Koyama, 2005) explained that "[d]isequilibria alone force the subject to go beyond his current state and strike out in new directions" (p. 10). Moreover, Waxer and Morton (2012) state that disequilibrium "motivates an individual to resolve the conflict and attain a new state of equilibrium" (p. 586). In Zazkis and Chernoff's (2006) terms, the cognitive conflict approach "allows students to trouble their own thinking, and it is through this conflict that they develop their own meanings, or at least seek to rectify the conflict" (p. 466). However, Piaget did not speak of a definite state of equilibrium. Instead, for Piaget (1977),

> equilibration is the search for a better and better equilibrium in the sense of an extended field, in the sense of an increase in the number of possible compositions, and in the sense of a growth in coherence. (p. 12)

In addition, whether an individual is aware of the conflict or not is another factor to consider. Balacheff (1986) claims that "[a] contradiction does not exist by itself but relative to someone who notices it. It has a witness. So, it may happen that it exists for one person and not for another one" (p. 10). Balacheff adds that noticing a contradiction depends on the individual's *background knowledge* and as such a need to analyze the individual's conceptions related to the problem to be solved emerges.

Piaget considers that both disequilibrium and cognitive conflict are crucial to cognitive development. Furthermore, in Piaget's theory, the state of equilibration or balance is important for cognitive growth. Campbell (2009) pointed out that "Piaget thought that development tend toward a balance, or equilibrium, between assimilation and accommodation" (p. 151).

Assimilation, Müller, Carpendale and Smith (2009) explain, "refers to the incorporation of new elements into already existing schemes… thereby giving meaning to those elements" (p. 4). Müller et al. provide the example of a "grasping scheme" and in grasping a new toy, the toy is assimilated to such scheme and acquires the meaning that it is an object that is "graspable". That means that individuals assimilate an object to a scheme by "extending scheme[s] to new objects" (Montangero & Maurice-Naville, 1997, as cited in Sellers, 2020, p. 56). In Swan's (2001) terms, assimilation "refers to the absorption of new ideas" (p. 156). Piaget (1977) emphasizes that "[a]ssimilation is clearly not a matter of passively registering what is going on around us" (p. 5). He points out that,

> [t]he organism is sensitive to a given stimulus only when it possesses a certain competence… A subject is sensitive to a stimulus only when he possesses a scheme that permits the capacity for response, and this capacity for response supposes a scheme of assimilation. (p. 5)

This means that while a person assimilates that one object is graspable, another may not. It is possible that the second person does not possess yet the "grasping scheme" and as a result for her/him an external-stimulus object does not have a meaning in that context.

On the other hand, Müller et al. (2009) explains that accommodation "refers to the modification of existing schemes to take account of particular features of the new object or situation" (p. 4). It "refers to the modifications that the child's cognitive structure makes as a result of 'fitting' new ideas into an existing framework" (Swan, 2001, p. 156). Sellers (2020) adds that, according to Piaget, there might be cases where an individual cannot assimilate an object into an existing scheme. Two contradicting schemes may be triggered by the same experience (e.g., when solving a math task), which leads to a "perturbation", or an "obstacle to assimilation" (Piaget, 1985, as cited in Sellers, 2020, p. 59). This happens as a result of cognitive inconsistencies that the individual experiences. In such a case, Sellers explains that the individual may adapt, or accommodate, her/his schemes in order to assimilate the object.

> In these moments, students have opportunities to consider their own thinking, and thus, accommodations may be viewed as the "source of change" and the mechanism by which learning occurs. (Montangero & Maurice-Naville, 1997, as cited in Sellers, 2020, p. 59)

Campbell (2009) illustrated the case of an unsuccessful assimilation that prepared the way for accommodation. He referred to a child's "fly-swatting scheme". When the child applied this scheme to another housefly, Campbell called that process a "routine assimilation". However, if the child tried to apply it to a hornet, the result would most likely be unsuccessful assimilation given that "swatting a hornet results in getting stung". In this case, Campbell explained that, the child needs to modify her/his scheme to accommodate it to the environment.

> Accommodating the scheme to the environment may mean putting restrictions on it (e.g., use this fly-swatting scheme only with nonstinging insects) or differentiating it into one or more subschemes. Sometimes entirely new schemes

> will need to be constructed for successful accommodation to take place. (Campbell, 2009, p. 151)

Swan (2001) adds that "Piaget held the view that individuals need time to interact with their environment (including the social environment) in order to construct concepts through assimilation and accommodation" (p. 156).

Furthermore, Steffe (1991) considers two types of accommodation: functional and metamorphic. By *functional accommodation* he means "any modification of the scheme that occurs in the context of using it" (p. 183) and by *metamorphic accommodation* "any modification that occurs independently and involves a general reorganization of a scheme" (p. 187). Sellers (2020) provides a clear illustration of all these processes in play. It is the example of a girl and her experience with air and motion. In her mind, the child finds connections between the movement of objects and air movement. She waves her hand in front of her face and she feels air. A new scheme, "motion-air", was just formed in her mind. In a windy day, the same girl sees trees moving. She assimilates the new situation by assuming that the new object, trees, is causing the air to move. On the beach, she sees no trees, but she notices that the air is blowing. She concludes that the waves are causing the air to move. She did not change her "motion-air" scheme. Instead, she *functionally accommodated* her scheme to use it given that there were no trees around causing the air to blow. She assimilated a new object (waves/ocean) into her "motion-air" scheme, but she did not reorganize her scheme. Now, suppose that the same girl sees a fan that is turned on and a piece of paper on the floor is moving across the room because of it. She notices that the fan is causing the air to blow, but also the air is causing the piece of paper to move. The girl begins to pay attention to other similar experiences. She blows on feathers and watches them move. She begins to distinguish when objects cause air movement and when the air causes objects to move. She reorganizes her mental "motion-air" scheme in such a way that both cases are considered and thus a *metamorphic accommodation* occurred in her mind. Possibly, she separated them into two different sub-schemes to assimilate those different experiences.

A contradiction with existing assumptions often triggers a disequilibrium and as Balacheff notices, it is usually regarded as an error; however, Brousseau asserted that,

> an error is not only the consequence of ignorance, or of uncertainty, or of an accident… An error may be the consequence of previous knowledge which had its own interest, its own successes, but which at the moment appears to be false, or more simply, not yet adapted. (Brousseau, 1976, as cited in Balacheff, 1986, p. 10-11)

In that respect, instead of using the terms "*mis*conceptions" or "*mis*assumptions", I opted for the general term *assumptions*. Indeed, I use the term *assumptions* in a broad sense. It may include implicit and explicit, mathematical and non-mathematical beliefs that individuals elaborate about meanings, inferences, properties, etc.

The assumptions that individuals use/hold are directly related to their background knowledge and past experiences. In that sense, they deserve close attention and should not be regarded as incorrect or correct through a simplistic contrast with expectations. For example, a student who only learned about natural numbers usually holds the assumption that the sum of two or more addends is bigger than any of the addends; while a student who learned about integers might hold a different but related assumption (e.g., the sum of two or more addends is not necessarily bigger than any of the addends). In a way, what might be regarded as an error is actually an assumption that worked properly well within a certain context or framework. If the context or framework changes, the assumption may experience some changes as well. Assumptions evolve with time. The assumptions we held when we were younger most likely differ from the ones we hold as grown-ups.

Notably, the assumptions that individuals hold in mathematics may be directly influenced by assumptions they use from other fields, like everyday life. For example, this is usual for the case of some quantifiers (e.g., see the case of the quantifier "some" in Chapter 2, Section I.6).

1.1. Structure of the Cognitive Conflict Approach

The main structure of the cognitive conflict approach is simply put into words in Limón (2001).

> (a) identifying students' current state of knowledge, (b) confronting students with contradictory information which is usually presented through texts… and interviewers, who make explicit the contradiction or only guide the debate with the student or among peers (small groups or whole classroom) …, or by the teacher and new technologies, and (c) evaluating the degree of change between students' prior ideas or beliefs and a post-test measure after the instructional intervention. (p. 360)

The approach does not only involve identifying the individuals' existing assumptions, but it also entails drawing their attention to those existing assumptions, notably, the ones that are expected to be challenged with the conflicting/contradictory cases. This has the purpose of making an explicit contrast with the conflicting case and so it does not pass unnoticed when it is brought forward. In this respect, Zazkis and Chernoff (2008) noticed that "inconsistency of ideas presents a *potential conflict*, it becomes a cognitive conflict only when explicitly invoked, usually in an instructional situation" (p. 196).

There may be different ways to make conflicts explicit. G. J. Stylianides and A. J. Stylianides (2009; 2014) consider that *conceptual awareness pillars* can be utilized with this goal.

> The notion of conceptual awareness pillars (or simply pillars) describes instructional activities that aim to direct students' attention to particular issues or to students' conceptions (understandings, beliefs, views, etc.) about those issues. (G. J. Stylianides & A. J. Stylianides, 2014, p. 385)

In terms of the contradictory or conflicting example, there might be situations where it may not trigger a conflict as expected. Zazkis and Chernoff (2008) pointed out that,

> counterexamples may not be sufficient in a conflict resolution, which is an ultimate goal of instruction. As teachers we are to seek strategic examples that will contribute not only to invoking a cognitive conflict but also to resolving it. (p. 197)

Zazkis and Chernoff (2008) introduced the notions of *pivotal example* and *bridging example*. A *pivotal* example, they explain, is a conflicting example that "creates a turning point in the learner's cognitive perception or in his or her problem solving approaches" and "may introduce a conflict or may resolve it". If the pivotal example supports the conflict resolution, then the authors call it *pivotal-bridging example* or just *bridging example*. A bridging example is "an example that serves as a bridge from learner's initial (naïve, incorrect or incomplete conceptions) towards appropriate mathematical conceptions" (p. 197).

Moreover, Zazkis and Chernoff (2008) used Watson and Mason's (2005) notion of *example spaces* and their classification to suggest what a conflicting example needs to satisfy to serve as a pivotal example.

> [A]n example should fit within, but push the boundaries of personal potential example spaces. That is, this should be an example that a learner accepts as "exemplary," but usually falls outside of personal/situated (i.e., immediately available and easy accessible) example space. (Zazkis & Chernoff, 2008, p. 206)

Zazkis and Chernoff (2008) also pointed out that the notion of a pivotal and bridging example is learner and situation dependent, which implies that a conflicting example may be pivotal/bridging for one learner, but not for another, or it can support her/his understanding in one situation, but not in another.

G. J. Stylianides and A. J. Stylianides (2009), unlike Zazkis and Chernoff, used the expression *pivotal counterexamples* to mean pivotal examples that introduce a conflict; that is, they did not include the case of examples that may support the conflict resolution.

1.2. Types of responses when facing contradicting information

To determine the effects that using the cognitive conflicts strategy may have, one could pay attention to the kind of responses individuals give. For instance, Limón (2001) described Piaget's considerations of *unadapted responses* and *adapted responses* when individuals were exposed to contradictory information.

> Unadapted responses are those where individuals do not realise the conflict. Adapted responses are classified into three types: *alpha, beta* and *gamma*. Alpha answers involve individuals who ignore or do not take into account the conflicting data. (p. 359)

Boom (2009) explains that, "[a]lpha reactions are characterized by the absence of any attempt to integrate the perturbations into the system in question" (p. 141). Limón (2001) continues to explain that *beta* responses are "characterised by producing partial modifications in the learner's theory, through generalisation and differentiation (generating an "ad hoc" explanation)" and *gamma* responses "involve the modification of the central core of the theory". Similar to Piaget's "unadapted responses", other studies have observed that individuals may not notice a contradiction (e.g., Tirosh & Graeber, 1990).

Chinn and Brewer (1993) suggested that there were seven possible responses to anomalous data: ignoring the anomalous data; rejecting the data; excluding the data; holding the data in abeyance; reinterpreting the data; peripheral changes and change of theory. While *ignoring the anomalous data* does not involve that the individual tries to explain the data, in *rejecting the data* "the individual can articulate an explanation for why the data should be rejected" (p. 6). In *excluding the data*, the individual considers that the anomalous data is outside the domain of the theory and as a consequence no change of the initial theory is achieved.

> In contrast to the individual who rejects anomalous data, the individual who excludes data from his or her theory does not have to make a judgement about the validity of the data. (p. 8)

For the case of *holding the data in abeyance*, the individual does not have the need to immediately explain the anomalous data; instead, s/he leaves the initial theory without changing it, as s/he assumes that someday in the future the initial theory will explain the data. Like in the previous response, *reinterpreting anomalous data* does not involve changing the initial theory. In this case the individual accepts the anomalous data as something that her/his initial theory could explain, by reinterpreting the data. *Peripheral change* of the initial theory entails minor modifications in the existing theory; however, the changes leave the main hypotheses of the initial theory unchanged. Even though this is the first response to anomalous data that involves some change in the initial theory, the individual does not really give up the initial theory. In contrast, *change of theory* mean that the individual need to change one or more core assumptions from the initial theory. In that case, the individual accepts the anomalous data and changes her/his existing theory to make sense of it.

Chan, Burtis, and Bereiter (1997) pointed out two general and contrasting approaches to understanding new concepts in unfamiliar domains: direct assimilation and knowledge building. *Direct assimilation* "involves fitting new information directly into existing knowledge". *Knowledge building* "involves learners treating new concepts as something problematic that need to explain" (p. 3). Chan et al. (1997) scaled these types of responses to new contradictory information into five levels, as follows:

> Level 1 (*Subassimilation*): the individual reacts to the contradictory information at an associative level. It is remotely related to the contradictory information. Her/his response does not deal with the main contradictory information, it is irrelevant to the point of the discussion (or text).
>
> Level 2 (*Direct Assimilation*): the individual assimilates the contradictory information as if it was something already known or excludes it if it does not fit with prior beliefs. Three types of responses within this level were identified: (a) *stonewalling*, where the individual ignores, denies, and excludes new information that differs from his or her beliefs; (b) *distortion*, where the individual distorts, twists and overinterprets the contradictory information to make it fit with prior beliefs; (c) *patching*, where the individual notices surface discrepancy and patches the differences by ad hoc rationalizations.
>
> Level 3 (*Surface-constructive*): the individual comprehends the contradictory information, but s/he does not consider its implications for her/his beliefs. Three types of responses are here distinguished: (a) *paraphrases*, where the individual paraphrases, makes simple local inferences from the new information, and asks related questions with no attempt to make belief revision; (b) *juxtaposition*, where the individual attends to the contradictory information but places new ideas alongside existing naive ideas with no integration, s/he juxtaposes correct and incorrect information as a way to minimize discrepancies between her/his beliefs and the new information; (c) *exception*, where the individual attends to the contradictory information but it is considered as an exceptional case with no need for belief revision.
>
> Level 4 (*Implicit knowledge building*): the individual treats the contradictory information as something problematic that needs explaining. It involves problem-solving procedures such as identifying inconsistencies, sense-making tactics, and construction of explanations. Two types of these responses were identified: (a) *problem recognition*, where the individual identifies the conflict and the contradictory information is recognized as something different from her/his beliefs and instead of eliminating the conflict or juxtaposing new and old information, the individual may treat her/his knowledge as an object of inquiry;

(b) *explanation-driven inquiry*, where the individual identifies inconsistencies and constructs explanations to reconcile knowledge conflict.

Level 5 (*Explicit knowledge building*): the individual accumulates the contradictory information for constructing coherence in domain understanding. This response involves addressing the problem of discordant pieces of information in an attempt to construct more complex knowledge. Two types of these responses are distinguished: (a) *coherence construction*, where the individual stops from making immediate judgement and examines connections among different pieces of information; (b) *comparison of conflicting models*, where the individual identifies conflicting hypotheses for explaining the domain in question.

(paraphrased from Chan et al., 1997, pp. 12–16)

As I see it, there are some overlaps between the types of responses I included in this section. In Table 1 I include the types of responses to conflicting information I organized based on the categories suggested by the authors I mentioned before. They can be first distinguished according to whether the individual realizes the conflict or not (first column of Table 1). A lack of realization of the conflict most likely will lead to no change at all. If the individual realizes the conflict, s/he can still decide to ignore/set aside the conflicting information and move on. S/he can also consciously change her/his initial assumptions. It can be a partial or a major change that takes place in the whole current system of assumptions aiming at consistency. In the second column of Table 1 I include the types of responses as I identified from the three references I made in this section. There I point the authors' specific type of response that resembles the description I gave. In the third column I focus on whether or not the type of responses leads to changing or not of the individual's initial theory as a result of facing conflicting information.

2. Proof, Examples and Language

Within this section I consider three parts. Because my focus is on proof-related issues, the first part includes my use of the term proof and some considerations about it that guided my research (Section 2.1). As I showed in Chapter 2, examples play a crucial role and have different status when proving is involved. Hence, the second part of this section has a focus on examples, a classification of examples that was notably important to my design of the intervention (Section 2.2). Last but not least, the third part has a focus on language, which I showed in Chapter 2 that is an important factor when dis/proving is involved (Section 2.3).

Table 1. Types of responses to conflicting data based on Chinn & Brewer's (1993) and Chan et al.'s (1997) research

Realization of the conflict	**Type of response to the conflicting data**		**Change in the individual's initial theory**
The individual does not realize/understand the conflict	The individual's response does not relate to the conflicting data	Similar to "unadapted response" (Piaget) and "subassimilation" (Chan et al., 1997)	NO change
	The individual *assimilates* the conflicting data to her/his existing schemes as if it was not conflicting with them		
The individual realizes/ understands the conflict	The individual *ignores* the conflicting data, without explaining why s/he ignores it	Similar to "ignoring anomalous data" (Chinn & Brewer, 1993) and "stonewalling" (Chan et al., 1997)	NO change
	The individual *rejects* the conflicting data. S/he believes the conflicting data is incorrect even though s/he may not be able to show that.	Similar to "rejecting anomalous data" (Chinn & Brewer, 1993) and "stonewalling" (Chan et al., 1997)	NO change
	The individual *excludes* the conflicting data. S/he does not argue whether the conflicting data is correct or not, it is merely excluded, possibly as an exceptional case.	Similar to "excluding anomalous data" (Chinn & Brewer, 1993), "stonewalling" and "exception" (Chan et al., 1997)	NO change
	The individual *distorts*, *overinterprets*, or *twists* the conflicting data to make it fit her/his initial theory, resorting to ad hoc explanations.	Similar to "reinterpreting anomalous data" (Chinn & Brewer, 1993), "distortion", "patching" and "juxtaposition" (Chan et al., 1997)	NO change
	The individual *postpones* the conflicting data, promising to deal with it later. S/he assumes that her/his initial theory will someday be formulated in such a way that it will explain the conflicting data.	Similar to "holding anomalous data in abeyance" (Chinn & Brewer, 1993)	NO change
	The individual *explores* the conflicting data, showing her/his (interest in further) understanding of it; however, s/he does not review her/his initial theory according to the new data.	Similar to "paraphrases" and "implicit knowledge building" (Chan et al., 1997)	NO change
	The individual makes *minor changes* to her/his initial theory. It does not involve changing the core assumptions of the initial theory. S/he changes her/his initial theory only at a local level.	Similar to "beta response" (Piaget) and "peripheral theory change" (Chinn & Brewer, 1993)	Partial change
	The individual makes *major changes* to her/his initial theory. It involves changing the core assumption(s) of the initial theory.	Similar to "gamma response" (Piaget), "theory change" (Chinn & Brewer, 1993), and "explicit knowledge building" (Chan et al., 1997)	Initial Theory change

2.1. Proof

In the field of Mathematics Education there is no consensus in what makes an argument a proof (e.g., Balacheff, 2008; Reid & Knipping, 2010; A. J. Stylianides, 2007; Czocher & Weber, 2020). Here I adopt the definition given by A. J. Stylianides (2007), which in my view is a suitable conceptualization since it is contextualized in the classroom community and it is flexible enough to be consistent with the perspective of the mathematical community.

> Proof is a mathematical argument, a connected sequence of assertions for or against a mathematical claim, with the following characteristics:
>
> 1. It uses statements accepted by the classroom community (set of accepted statements) that are true and available without further justification;
>
> 2. It employs forms of reasoning (modes of argumentation) that are valid and known to, or within the conceptual reach of, the classroom community; and
>
> 3. It is communicated with forms of expression (modes of argument representation) that are appropriate and known to, or within the conceptual reach of, the classroom community. (p. 291)

While the first component includes, for example, axioms, definitions, theorems that have been accepted as starting points to prove, G. J. Stylianides and A. J. Stylianides (2008) pointed out that the valid mode of argumentation implied in proving should be deductive as that is the key form of reasoning for necessary inferences.

This does not mean that other forms of reasoning are not important. Indeed, they are; for example, when coming up with conjectures and possibly on the way to find a dis/proof. However, a deductive mode of argumentation is what provides conclusiveness to a mathematical proof. This form of (deductive) reasoning is refined within the classroom community, which A. J. Stylianides (2007) refers to in his definition of proof. Examples of deductive modes of argumentation are proof by contradiction, proof by induction, generic proofs, valid rules of inference such as modus ponens and modus tollens, etc. In contrast, an *empirical argument* does not qualify as a proof. It is, as explained by A. J. Stylianides and G. J. Stylianides (2009), "an invalid argument that provides inconclusive evidence for the truth of a statement by verifying its truth in a proper subset of the cases in the domain of the statement" (p. 239). I would add that the case of empirical arguments does not include cases of generic examples that may qualify as proof. In addition, my view is that it is possible that a student provides an empirical argument, even though s/he might know that such an argument does not qualify as a proof, but s/he cannot manage to provide something better (as I explained in Chapter 2, Section I.2). In those cases, a distinction should be made between the individuals' assumptions about proof and the actual "proofs" they present.

As examples of the modes of argument representation involved in a proof, A. J. Stylianides (2007) includes: linguistic (e.g., oral language), physical, diagrammatic/pictorial, tabular, symbolic/algebraic, etc. G. J. Stylianides and A. J. Stylianides (2008) consider that no specific form of expression or mode of argument representation is necessarily associated with the deductive mode of argumentation. In that sense, a proof could be given by using diagrams, algebraic terminology, pictures, or in a narrative form. As long as the deductive mode of argumentation is present and the form of expression is accepted within the classroom community, there is no reason not to qualify an argument as a proof.

2.2. Examples

I use the term *example* to mean an instantiation of something in relation to a specific statement. Notably, within the category of "examples" I consider more specific subcategories, namely: *confirming examples*, *irrelevant examples*, and *counterexamples*.

Examples play different roles when dis/proving. In particular, during the 2018-intervention, *irrelevant* and *confirming* examples played an important role in the process of finding a general characterization of counterexamples to a UAS. I adopt Buchbinder and Zaslavsky's (2009) definition for those two cases.

An *irrelevant example* for the universal statement "for all x that belong to a particular domain D, the proposition $P(x)$ is true" as an example of an element that does not belong to the domain D. For example, irrelevant examples for the universal statement *"All even numbers are palindrome"* are 7, 11, 353 since they are not even numbers; that is, they do not belong to the *set of analysis* (see Section II.1 below) or *domain*, which in this case is determined by all even numbers.

A *confirming example* for the same statement ("for all x that belong to a particular domain D, the proposition $P(x)$ is true") is an example of an element x in the domain D for which $P(x)$ is true. For example, confirming examples for the same statement *"All even numbers are palindrome"* are 202, 45354 and 6336. They are confirming examples given that they are even numbers and palindrome; that is, they satisfy both conditions in the statement. Sometimes I use the expressions *supporting* or *verifying examples* as interchangeable with *confirming examples.*

A *counterexample* for the same statement "for all x that belong to a particular domain D, the proposition $P(x)$ is true" is an example of an element x in the domain D for which $P(x)$ is not true. For example, a counterexample for the statement *"All even numbers are palindrome"* is the number 4310, which is even but is not palindrome. Moreover, Peled and Zaslavsky (1997) distinguished three types of counterexamples according to its nature:

- *Specific*: A counter-example which satisfies the task, as it provides a specific example contradicting the claim, yet does not give a clue to an underlying mechanism for constructing other (similar or related) counter-examples.
- *Semi-General*: A counter-example which provides an idea about a mechanism for generating other (similar or related) counter-examples for the claim, yet does not tell "the whole story" and does not cover "the whole space" of counter-examples.
- *General*: A counter-example which provides a "behind the scene story," and suggests a way to generate the entire counter-example space. (p. 53)

These types of counterexamples are notably relevant for my analysis of the teachers' responses to our discussions about the disproving of universal statements.

In this context, language also plays an important role since the way individuals understand a specific term may differ from person to person. For instance, the way an individual uses the term *example* might be different from another person. Someone can use the term "example" to mean only a confirming example, while another person can have a broad usage of the term and accept irrelevant examples for the "examples" category. Paying attention to this issue must be considered. In the following section I attempt to do so.

2.3. Language

The way language is used in everyday life is usually different from the way it is used in mathematics. In Halliday's (1978) terms, "[l]anguage, unlike mathematics, is not clearcut or precise. It is a natural human creation, and, like many other natural human creations, it is inherently messy" (p. 203).

Difficulties may arise when formalizing terms that are already used in everyday language since, Halliday (1978) asserted, "a term for a mathematical concept may also exist as an element in natural language, and so carry with it the whole semantic load that this implies" (p. 203).

Mathematical Language

Halliday (1978) discussed the notion of "register" in general and the "mathematics register" in particular. He considered that a register is

> a set of meanings that is appropriate to a particular function of language, together with the words and structures which express these meanings. We can refer to a 'mathematics register', in the sense of the meanings that belong to the language of mathematics (the mathematical use of natural language, that is: not mathematics itself), and that a language must express if it is being used for mathematical purposes. (p. 195)

One important challenge is supporting students' transit from informal ways of using language towards technical/formal ones; in particular, the transit from the informal language to the mathematical. Schleppegrell (2007) pointed out that, "[l]earning the language of a new discipline is a part of learning the new discipline; in fact, the language and learning cannot be separated" (p. 140). This involves many challenges for many students (see Schleppegrell, 2007, for a review).

An important example of the differences in language is the way some quantifiers are used in the mathematical register and in the everyday-language register. For instance, the quantifier "some" in everyday language is used to mean "some, but not all"; whereas its meaning in mathematics does not deny the possibility of *all*: "some" means "some and maybe all" (e.g., Epp, 2003; Lee & Smith, 2009; Woodworth & Sells, 1935). This can trigger several difficulties in particular when "some-statements" are the focus in proof-related activities.

Given that the everyday-language register has an influence on other registers, it is important to understand the principles that rule this register of language.

Maxims of Conversation

In the context of colloquial language, Grice (1975; 1989) described the *cooperative principle* that governs conversations as, "[m]ake your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged." Grice (1989) distinguished four categories or maxims of conversation within the cooperative principle; namely:

> (1) *Quantity* relates to the quantity of information to be provided, and it encompasses the two submaxims:

(a) Make your contribution as informative as is required (for the current purposes of exchange);

(b) Do not make your contribution more informative than is required.

(2) *Quality*, under which falls the supermaxim “Try to make your contribution one that is true”, and the two submaxims:

(a) Do not say what you believe to be false;

(b) Do not say that for which you lack adequate evidence.

(3) *Relation*, that includes the only “Be relevant”.

(4) *Manner*, which is not directly related to what is said, but to *how* what is said is to be said. It includes the four submaxims:

(a) Avoid obscurity of expression;

(b) Avoid ambiguity;

(c) Be brief (avoid unnecessary prolixity);

(d) Be orderly. (p. 26–27)

It is important to have in mind these principles as they might explain several challenges individuals face when engaged in proof-related activities.

II. Mathematical Framework

When it comes to mathematical reasoning and proving, universal and existential statements play an important role. Most theorems in mathematics are universal statements, which are closely related to existential statements: the negation of a universal statement is an existential statement, and the other way around, the negation of an existential statement is a universal statement. Therefore, understanding what it means a universal statement to be false involves understanding what it means for its negation to be true, and vice versa. Thus, understanding what constitutes a conclusive argument to prove a universal statement is false involves understanding what conclusive evidence is necessary to show when proving that its negation is true.

In this section I include some elementary mathematical/logic-related concepts that are relevant to my research. I divide the chapter into three sections. Section 1 has a focus on the logical interpretation of single-quantified statements. Section 2 draws attention to the representation of single-quantified statements. Section 3 is related to proving and disproving of single-quantified statements.

1. Logical Interpretation of Single-Quantified Statements

Central to the design of my interventions and my analysis of the teachers’ understandings is the logical interpretation of single-quantified mathematical statements. In this section I explain what that entails.

The mathematical statements that were the focus of my interventions are what is known as *single-quantified statements* (SQ-statements). SQ-statements can be divided into two main groups: *universal* and *existential* statements.

When the statement applies to every element under consideration, it is a *universal statement* (henceforward US); whereas, when it applies to at least one element, it is an *existential statement* (henceforth ES). For instance, *"All numbers divisible by 5 are square numbers"* is a universal statement, whereas *"There exist prime numbers that are not perfect square numbers"* is an existential statement.

A mathematical statement can be either true or false. I refer to these two possible outcomes as "truth values".

Set of analysis in a SQ-statement

In the case of mathematics, as well as in common language, a quantified statement is commonly formulated considering a domain restriction (Abrusci, Pasquali & Retoré, 2016). Abrusci et al. used the statement *"All the stars go away"* to show that the sentence is restricted to the class of "stars". In my work with the teachers it was important to identify the class of elements to which the quantifier applied.

The *set of analysis* in a statement is the set the quantifier applies to, it includes all the elements that satisfy the first, *antecedent* condition, in the statement. For example, in the statements *"All numbers divisible by 5 are square numbers"* and *"There are numbers divisible by 5 that are square numbers"* the *set of analysis* is the same, namely, the set of all numbers divisible by 5. In particular, in the context of my work with primary school teachers, most of the mathematical statements we discussed were framed in the set of natural numbers; however, in general, that does not have to be the case.

The notion of *set of analysis* has a long history. Irving Copi called it *subject class*, other authors called it *domain* (e.g., Buchbinder & Zaslavsky, 2009; Buchbinder, Zodic, Ron & Cook, 2017) or *cases involved* (e.g., A. J. Stylianides & Ball, 2008). Additional ways I use to refer to it are *set of reference, antecedent set, set of cases involved in the statement, if-condition set,* etc.

Understanding the logical interpretation of a SQ-statement

I consider that understanding the logical interpretation of a SQ-statement mainly entails identifying and understanding the interwoven link between three components of statements: (1) the two sets involved in the statement: the set of analysis and the conclusion set (or "predicate class", as Copi, Cohen & McMahon, 2014, call it); (2) the elements from the set of analysis to which the statement applies (that is the quantification); and (3) the claim made in the statement about those elements.

For instance, in the statement, *"All numbers divisible by 5 are square numbers"*, the *set of analysis* is the set of all numbers divisible by 5 and the *conclusion set* is the set of all square numbers. The statement is a universal statement; it applies to every element in the set of analysis. Hence, its logical interpretation is that every number divisible by 5, without any exceptions, is a square number.

The claim made in a SQ-statement can be that all or some elements in the set of analysis belong to the conclusion set, as in *"All numbers divisible by 5 are square numbers"*, where the claim is that every element in the set of analysis belong to the conclusion set. On the other hand, the claim made in a SQ-statement can be that all or some elements in the set of analysis *do not* belong to the conclusion set, as in the universal statement *"All numbers divisible by 5 are not square numbers"*. The first kind of statements are *affirmative* statements, and the second kind are *negative* statements.

There are four possible combinations of the two quantifications (all/some) and the two possibilities affirmative/negative. I use the expression *Universal Affirmative Statements* (UAS) to refer to universal statements that are affirmative, for example the statement of the form *"All X are Y"*. Universal statements that are negative, for instance the statement of the form *"All X are not Y"*, I call *Universal Negative Statements* (UNS). I use the expression *Existential Affirmative Statement* (EAS) to refer to statements, for instance, of the form *"Some X are Y"*. And finally, I call statements of the form *"Some X are not Y"*, *Existential Negative Statements* (ENS).

Identifying the three aforementioned components in a statement leads to establishing a distinction between universal and existential statements, and their subcategories. Figure 1 (below) summarizes these distinctions.

Note that the logical interpretation of a SQ-statement is independent of its truth value. Even though *"All numbers divisible by 5 are square numbers"* is a false statement, it is possible to interpret it logically; that is, determining what the set of analysis is, whether the statement refers to all or some of the elements in the set of analysis, what is said about those elements, etc.

Negations

Negation is an important topic and this is particularly true in mathematics. Vender and Delfitto (2010) put it in the following terms,

> Negation is a highly specific linguistic tool, peculiar of human language, which is employed to accomplish different tasks such as denying, contradicting, refusing concepts, correcting wrongly made inferences, but also lying and speaking ironically.
>
> For its fundamental role in human language, negation has been extensively studied throughout the centuries. It has been a matter of research for philosophers as Plato and Aristotle and it has been dealt with in classical logic. (p. 2)

In concrete, the negation of a universal statement is an existential statement, and vice versa, the negation of an existential statement is a universal statement. Moreover, the negation of a UAS is an ENS; whereas the negation of a UNS is an EAS. The negation of an ENS is a UAS, while the negation of an EAS is a UNS (see Table 2).

Table 2. Type of SQ-statement and its respective negation. Exemplifications.

Type of SQ-statement	Negation	Example	Negation
Universal Affirmative Statement (UAS)	Existential Negative Statement (ENS)	*All numbers divisible by 5 are square numbers*	*Some numbers divisible by 5 are not square numbers*
Universal Negative Statement (UNS)	Existential Affirmative Statement (EAS)	*All numbers divisible by 5 are not square numbers*	*Some numbers divisible by 5 are square numbers*
Existential Affirmative Statement (EAS)	Universal Negative Statement (UNS)	*Some numbers divisible by 5 are square numbers*	*All numbers divisible by 5 are not square numbers*
Existential Negative Statement (ENS)	Universal Affirmative Statement (UAS)	*Some numbers divisible by 5 are not square numbers*	*All numbers divisible by 5 are square numbers*

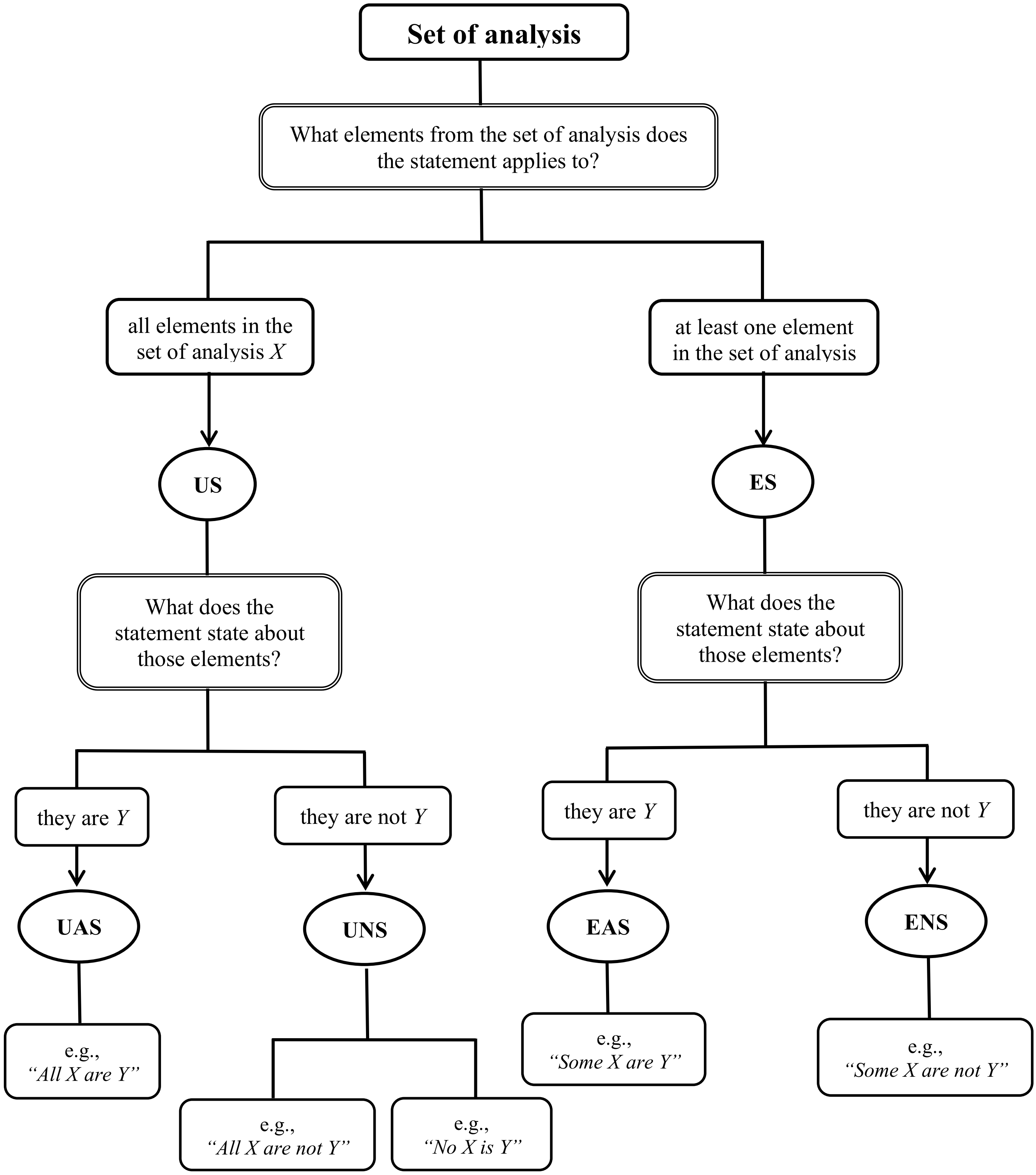


Figure 1. The logical interpretation of SQ-statements. Their classification.
(US: Universal Statements; ES: Existential Statements; UAS: Universal Affirmative Statements; UNS: Universal Negative Statements; EAS: Existential Affirmative Statements; ENS: Existential Negative Statements)

Horn (2001) explained that "Aristotle's theory of negation has its roots within his system of oppositions between pairs of terms" (p. 6) and he stated that there were four species of opposition: correlation (e.g., double vs. half), contrariety (e.g., good vs. bad; white vs. black), privation (e.g., blind vs. sighted) and contradiction (e.g., he sits vs. he does not sit). In mathematics, the negation of a mathematical statement is its contradictory statement, which does not only have its opposite truth value, but also follows these three laws:

> The Law of Contradiction: two opposing statements cannot both be true ("2 is even" and "2 is not even"),
>
> The Law of Excluded Middle: two opposing statements cannot both be false, and
>
> The Law of Double Negation: if the statement opposing a given claim is false, then the original statement must be true. (Dawkins, 2017, p. 503)

Figure 2 is a modification of the traditional square of opposition that includes the relation of SQ-statements by contrariety and contradiction (see e.g., Horn, 2001, p. 10).

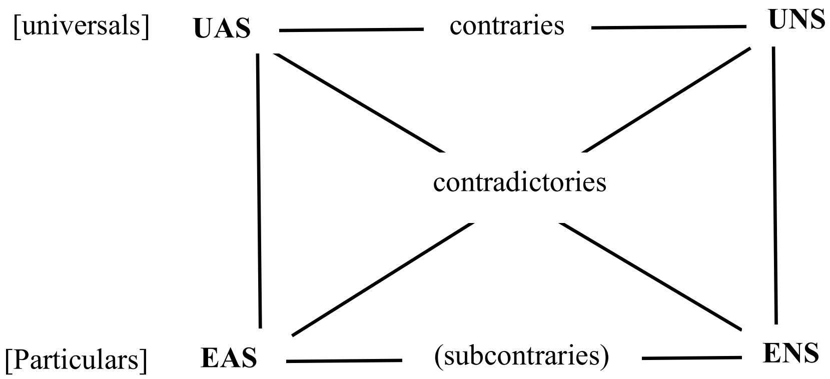


Figure 2. A modification of the traditional square of opposition for SQ-statements.

A common misconception about negations is to assume that the negation of the UAS *"All X are Y"* is the respective UNS *"No X are Y"* (Alacaci & Pasztor, 2005; Epp, 1999). In other terms, it is usual to presume that the negation of a UAS is its *contrary*, instead of its *contradictory* (see Figure 2).

2. Representations of SQ-statements

In everyday language there are many ways to express each type of SQ-statement as alternative forms of expression. There are also two forms of graphical representations of SQ-statements that I use in my interventions.

Representations of a UAS

A universal statement of the form *"All X are Y"* can be expressed in many equivalent ways. Among them:

- *Each element in X is an element in Y*,
- *If an element belongs to X, then it belongs to Y*,
- *For every x in X, x belongs to Y*,
- *The X are Y*,
- *Any X is Y*,
- *Each X is Y*,

- *Every X is Y.*

In mathematics, all of these are equivalent statements even though the universal quantifier involved in them is not exactly the same (see Copi et al., 2014, p. 443, for more equivalent forms). They are *equivalent* statements in the sense that they have the same logical interpretation. The universal quantifier can take different forms, such as: “for every”, “for each”, “for any”, “given any”, or “for all”. For example, the universal statement *“All dogs are playful”* can be also expressed as *“Every dog is playful”*, *“Any dog is playful”*, *“If it is a dog, then it is playful”*, *“Dogs are playful”*. If we let X be the set of all dogs, then we can express the original statement as *“For all k in X, k is playful”*.

Within the category of universal statements, I consider sub-categories depending on the the universal quantifier involved. I refer to them as “all-statements”, “every-statements”, “any-statements”, “each-statements”, etc. For instance, the statement *“All dogs are playful”* is an example of an “all-statement”, while *“Each dog is playful”* is an “each-statement”.

A visual representation for the logical interpretation of the UAS *“All X are Y”* could be given by using a Venn diagram (as shown in Figure 3) or a Euler diagram (as shown in Figure 4).

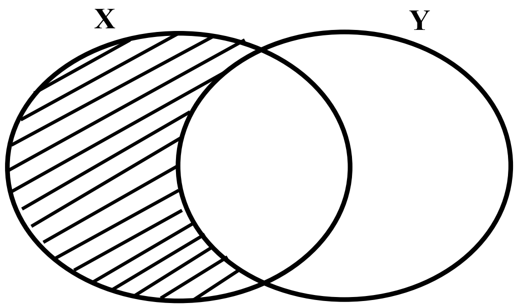


Figure 3. Venn diagram for the logical interpretation of the universal affirmative statement *“All X are Y”*

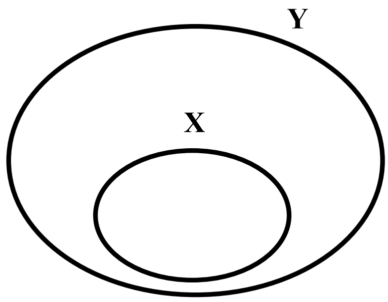


Figure 4. Euler diagram for the logical interpretation of the universal affirmative statement *“All X are Y”*

In Figure 3, set $X - Y$ has been shaded to represent that it is empty. In both cases, it is represented that every element in the set of analysis X belongs to the conclusion set Y, or in other words, set X is included in set Y.

Universal Conditional Statements

An important form of UASs is the conditional (if-then) form. In simple terms, a *Universal Conditional Statement*[9] (henceforth UCS) is a statement of the form *“If X, then Y”* that is equivalent to a universal statement, e.g. *“All X are Y”*. X is called the *antecedent* and Y is the *consequent* of the conditional statement. For example, the statement *“If a natural number is even, then it is a square number”* is a UCS and can be rewritten as *“For every natural number x, if x is even, then x is a square number”* or *“All even natural numbers are square numbers”*.

Most mathematical statements are universal conditional statements (Durand-Guerrier, 2003), with the form, “[f]or all elements x in a certain set, if <hypothesis> then <conclusion>” (Epp, 2009b, p. 314). Epp (2020, p. 113) claimed that this is the most

[9] Sometimes I use the expression “if-then-statements” to mean conditional statements.

important form of statement in mathematics. Additionally, given that my focus is on mathematics, my use of the expression *conditional statements* (CSs) involves all those statements where the antecedent is assumed to be true; that is, if the CS *"If A, then B"* is under discussion, I start to reason from the assumption that I count on *A* and I need to find out what inferences would lead me to conclude *B*. In that sense, my analysis of the *"if-then"* structure relies on the set made up of elements that satisfy the antecedent, *"if-condition"*, or the *set of analysis*, since that is basically the assumption I begin my analysis from in order to develop a chain of inferences that would potentially lead me to conclude the consequent.

Representations of a UNS

A universal negative statement *"All X are not Y"* has these equivalent forms, among others:

- *Each element in X is not an element in Y,*
- *If an element belongs to X, it does not belong to Y*,
- *For every x in X, x does not belong to Y,*
- *No X is Y*,
- *The Xs are not Y.*

A visual representation for the logical interpretation of the UNS *"All X are not Y"* is given in Figure 5, by using a Venn diagram, or in Figure 6, with a Euler diagram.

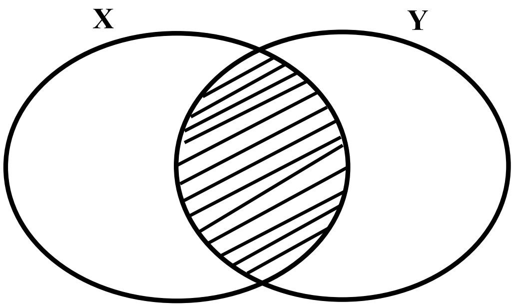


Figure 5. Venn diagram for the logical interpretation of the universal negative statement *"All X are not Y"*

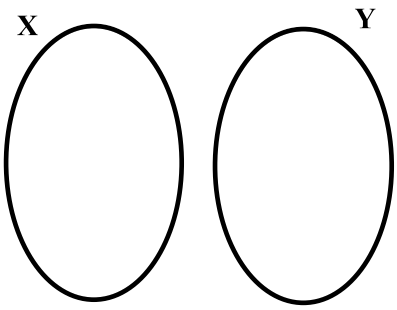


Figure 6. Euler diagram for the logical interpretation of the universal negative statement *"All X are not Y"*

In Figure 5, set $X \cap Y$ has been shaded to represent that it is empty. In both cases, it is represented that every element in the set of analysis *X* does not belong to the conclusion set *Y*, or in other words, set *X* is completely excluded from set *Y*, or that both sets (*X* and *Y*) are disjoint.

No-statements

A special form of a universal negative statement I call "no-statements" (henceforth NO-Ss). These are statements of the form *"No X is Y"* (affirmative NO-S) as well as *"No X is not Y"* (negative NO-S). They are universal statements because they apply to all elements in *X*. Notably, they are equivalent to *"All X are not Y"* and *"All X are Y"*, respectively. For example, the statement *"No dog is playful"* can be rewritten as *"All dogs are not playful"* without changing its logical interpretation. In both cases, the statements refer to "every dog" or "all dogs". An interesting example of NO-S is the negative NO-S *"No rhombi are not parallelograms"* (example taken from Dawkins, 2017, p. 512), which is equivalent to *"All rhombi are parallelograms"*.

From now on I do not make any distinction among the statements *"None of the X is Y"* and *"No X is Y"*.

Representations of an EAS

An existential affirmative statement *"Some X are Y"* is equivalent to the following statements:

- *At least one element in X is an element in Y*,
- *For at least one x in X, x belongs to Y.*
- *There is X that is Y*,
- *There exists X that is Y*
- *At least one X is Y.*

The existential quantifier can take the form "for some", "for at least one", "there exists", "there is", or "there is at least one". For example, *"There are dogs that are playful"* is an existential statement, which can also be expressed as *"Some dogs are playful"*. If we let *D* be the set of all dogs, then we can express the original statement as *"There exists x in D such that x is playful"*.

Similar to the case of universal statements, within the category of existential statements, I consider sub-categories to which I refer as "there-exist-statements", "some-statements", "there-is/are-statements", "at-least-one-statements", etc., depending on the form of the existential quantifier involved. For instance, the statement *"There are dogs that are playful"* is an example of an "there-are-statement" and *"Some dogs are playful"* is a "some-statement".

A visual representation for the logical interpretation of the EAS *"Some X are Y"* is given by using the Venn diagram in Figure 7.

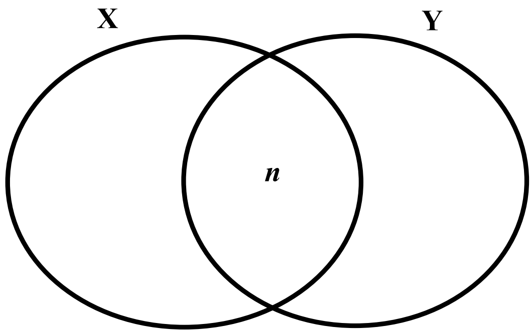


Figure 7. Venn diagram for the logical interpretation of the existential affirmative statement *"Some X are Y"*

In Figure 7, "*n*" represents the existence of at least one element in both sets, *X* and *Y*, or in other words, that the intersection $X \cap Y$ is non-empty. It is represented that at least one element in the set of analysis *X* belongs to the conclusion set *Y*, or in other words, set *X* and set *Y* have at least one element in common.

Representations of an ENS

The existential negative statement *"Some X are not Y"* is equivalent to:

- *There are X that are not Y*,
- *At least one element in X is not an element in Y*,
- *There exists one element that belongs to X and it does not belong to Y*,
- *For at least one x in X, x does not belong to Y.*

For instance, the statement *"Some dogs are not playful"* is an ENS, which might also be expressed as *"At least one dog is not playful"*.

A visual representation for the logical interpretation of the ENS *"Some X are not Y"* is given in the Venn diagram in Figure 8.

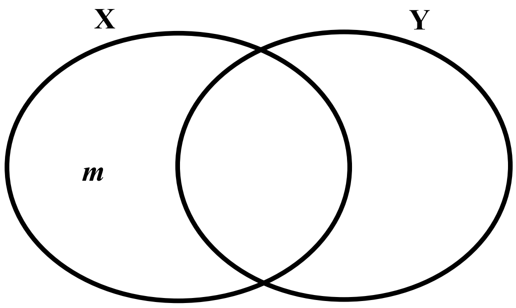


Figure 8. Venn diagram for the logical interpretation of the existential negative statement *"Some X are not Y"*

In Figure 8, "*m*" represents the existence of at least one element in *X* that is not an element of *Y*, or in other words, that set $X - Y$ is non-empty. In Figure 8 it is represented that at least one element in the set of analysis *X* does not belong to the conclusion set *Y*.

Unlike the case of making sound valid inferences, in order to find equivalent statements (statements with the same logical interpretation), or represent a statement visually, true statements are not necessarily required as starting points (Baggini & Fosl, 2010).

Representations of Negations

A negation can be expressed in different ways. As illustrations, the negation of the universal affirmative statement *"All X are Y"* can be also expressed as *"It is false that all X are Y"*, *"Not all X are Y"*, *"It is not the case that all X are Y"*, etc. Similarly, the negation of an existential negative statement of the form *"There exists X that is not Y"* can be expressed as *"It is not true that there exits X that is not Y"*, *"There does not exist X that is not Y"*, etc.

3. Proving and disproving of SQ-statements

The dis/proof of SQ-statements is determined by the quantifiers involved in them. Epp (2020) also highlights this and calls attention to cases where quantifiers are left implicit.

> The quantification of a statement – whether universal or existential – crucially determines both how the statement can be applied and what method must be used to establish its truth. Thus it is important to be alert to the presence of hidden quantifiers when you read mathematics so that you will interpret statements in a logically correct way. (p. 117)

Examples of SQ-statements where the quantifier is implicit are *"Rectangles are squares"*, which refers to all rectangles (i.e., it is equivalent to *"All rectangles are squares"*). *"The* *number 25 can be written as the square power of a natural number"* which can be rephrased as the existential statement *"There exists a natural number such that its square power is 25"*.

In simple terms, the UAS *"All X are Y"* is true if each element in *X* is an element in *Y* (all elements in *X* are included in *Y*). It is false, if there is at least one element in *X* that does not belong to *Y* (a counterexample). In that respect, a counterexample for the statement *"All X are Y"* must satisfy condition *X*, but must not satisfy condition *Y*.

There are different proof methods that show that a universal statement is true. Among them, I focus here on two of them: direct proof and proof by exhaustion. Epp (2009b, p. 314) explains that a direct proof for the statement "For all elements *x* in a certain set, if <*hypothesis*> then <*conclusion*>" has the following outline:

- Suppose that *x* is a particular but arbitrarily chosen (or "generic") element of the set for which the hypothesis is true.
- We must show that *x* also makes the conclusion true.

For example, the universal statement *"All numbers[10] divisible by 6 are even numbers"* is true and can be proved with a direct proof as follows:

> Let x be a number divisible by 6. This means that x has the form $x = 6k$ for some $k \in N$. It can be rewritten as $x = 2(3k)$, which shows that x is even, given that x is divisible by 2; that is, x has the form $x = 2m$ for some natural number m (here, $3k$). Hence, x is an even number.

An example of proof by exhaustion is the following that shows that the statement *"All one-digit natural numbers[11] divisible by 6 are even"* is true.

> The only one-digit natural numbers that are divisible by 6 are 0 and 6 and they are even.

A proof by exhaustion consists of verifying that every case in the set of analysis in a universal statement meets the conclusion condition.

On the other hand, the universal statement *"All square numbers are prime numbers"* is false because 36 is a square number; however, it is not a prime number. This means that 36 is a counterexample to the statement.

Further, disproving/refuting a UAS of the form *"All X are Y"* (or proving that the UAS is false) is the same as proving its negation, the ENS *"Some X are not Y"* (Abrusci et al., 2016, p. 195; Epp, 2020, p. 165). Indeed, visually, if it is not true that *"All X are Y"*, this means that not all elements in *X* are elements in *Y*; therefore, there must be elements in *X* that are not elements in *Y*. In other terms, if the diagram in Figure 3 (see above) is false, it implies that set $X - Y$ should not be empty and hence it must have at least one element in it (see Figure 9 below), which is guarantee or proof that *"Some X are not Y"* is true, which is the negation of the original universal statement.

Likewise, the existential statement *"There exists X that is Y"* is true if there is at least one element in *X* that belongs to *Y*. The ES is false if none of the elements in *X* belong to *Y*; that is, its negation, *"All X are not Y"* or *"No X is Y"*, is true.

[10] In the direct proof I suggest here I considered the set of natural numbers as my *universal set* because that was the focus with the teachers in my interventions; however, in other contexts the universal set is the set of integers.

[11] Herein I consider number zero (0) as a natural number.

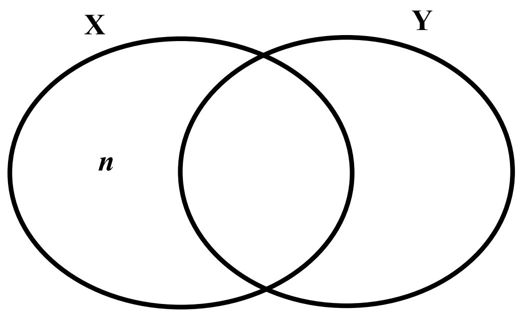


Figure 9. Visual representation that the statement *"All X are Y"* is false.

For example, the ES *"Some numbers divisible by 6 are even numbers"* is true because there is at least one number divisible by 6 that is even, say 24. The ES *"There exist numbers divisible by 6 that are not even"* is false because, as I proved before, all numbers divisible by 6 are even.

Non-minimal justifications

The justifications I gave in the previous lines are what I call "sufficient dis/proofs", or they are also called "minimal dis/proofs". Nonetheless, there are also other justifications that may qualify as valid (dis/proofs) and are *non-minimal justifications*. For example, a non-minimal proof that the ES *"Some numbers divisible by 6 are even numbers"* is true may be a proof that all numbers divisible by 6 are even (such as the direct proof I gave above), or a proof that all numbers divisible by 6 smaller than 100 are even. Similarly, a non-minimal disproof for the US *"All square numbers are prime numbers"* may be a proof that shows that no square number is a prime number.

Two types of counterexamples suggested by Peled and Zaslavsky (1997) fit within the category of non-minimal disproofs; namely: *semi-general* and *general* counterexamples (see Section I.2.2 above).

Chapter 4: Methodology and Design

"The first and most compelling argument for initiating design research stems from the desire to increase the relevance of research for educational policy and practice."

(van den Akker, Gravemeijer, McKenney & Nieveen, 2006, p. 3)

In this chapter I discuss the methodology and the design I used to develop my research. Design-based research (henceforth, DBR) is a methodology that matches the focus of my research. I aim at developing design principles for future teaching development interventions with a focus on proof and explaining how they can be used with this goal. The idea of improving the design principles is implicit in the cyclic nature of the Design Based Research methodology. DBR contributes with the generation of different theory elements (see Prediger, 2019). Notably, I pursue the identification of explanatory and predictive theory elements in relation to my research questions (see Chapter 1). In that process, categorial theory elements may also emerge as a result of my attempts to explain the observed phenomena.

I. Design-Based Research

Design-Based Research is a research methodology that links theory and practice in educational settings (Goff & Getenet, 2017). Bakker and van Erde (2015) pointed out that DBR

> aims both at developing theories about domain specific learning and the means that are designed to support that learning. DBR thus produces both useful products (e.g., educational materials) and accompanying scientific insights into how these products can be used in education. (p. 430)

Cobb et al. (2003) identified five characteristics of DBR that were summarized in Bakker and van Erde (2015, p. 437) as: (1) its purpose is to develop theories about learning and the means that are designed to support that learning. (2) It has an interventionist nature, where "learning already takes place in learning ecologies as they occur in schools and thus methods measure better what researchers want to measure, that is learning in natural situations" (p. 438). (3) It has prospective and reflective components that need not be separated by a teaching experiment as "[r]eflection can be done after each lesson, even if the teaching experiment is longer than one lesson" and as a result of that reflection changes in the original design can be implemented for the next lesson. (4) It has a cyclic nature, with the cycles usually consisting of three phases: preparation and design, teaching experiment, and retrospective analysis. (5) The theory under development has to do real work, which means that even though the theory is developed for a specific domain, it needs to be able to be transferred to different contexts (e.g., a different classroom, a different school, another country).

In relation to the phases that are part of a cycle of DBR, the first phase, *preparation and design*, includes the identification of a problem of interest, the review of the related literature, the development of an initial theoretical framework for the design of the intervention and the design of tasks that are part of the intervention and have specific goals.

The second phase is not exactly identical to what Steffe et al. (2000) call *teaching experiment*, but it is similar to it. It focuses on the implementation of the designed intervention. It aims at "understanding the progress students make over extended periods"

(Steffe et al., 2000, p. 273). In this phase the initial hypotheses are tested. During this period, though,

> researchers do their best to "forget" these hypotheses during the course of the teaching episodes in favor of adapting to the constraints they experience in interacting with the students. (Steffe et al., 2000, p. 275)

New hypotheses might emerge during the teaching episodes and the teacher-researcher may formulate and test those hypotheses in a following teaching episode. The researcher should focus on what happens during the teaching episodes and,

> [r]ather than believing that a student is absolutely wrong or that the student's knowledge is immature or irrational, the teacher-researcher must attempt to understand what the student can do; that is, the teacher-researcher must construct a frame of reference in which what the student can do seems rational. (Steffe et al., 2000, p. 277)

In that process the researcher can modify the goal structure according to the students' mathematical activity. Steffe et al. (2000) explain that "[e]xtending and modifying the goal structure lasts until the students' schemes seem to be well established and the students seem to have reached a plateau" (p. 280).

Data collection during a teaching experiment includes video/audio recordings of all lessons, video/audio recordings of possible interviews, students' work, tests before and after instruction, and field notes (Bakker & van Erde, 2015).

Later, the researcher returns to the research hypotheses, after the teaching episodes are finished, during the last phase of DBR, *the retrospective analysis*. This last phase is very important and it usually involves more work than the teaching experiment itself. A retrospective analysis entails a careful analysis of videotapes, which activates the interaction the researcher had with the students and bring those memories into the researcher's conscious awareness (Steffe et al., 2000). In this process the research can engage in a reconstruction of the students' mathematical activity. In this phase the researcher can take the time to carefully reflect on the interpretations that s/he could have quickly made while the teaching experiment took place.

> [T]hrough watching the video-tapes, the teacher-researcher has the advantage of making an historical analysis of the students' mathematics retrospectively and prospectively, and both of these perspectives provide insight into the students' actions and interactions that were not available to the teacher-researcher when the interactions took place. It is especially important that the teacher-researcher be able to take a prospective view of the interacting child and interpret the significance of what the students may or may not have been doing from that perspective. In this way, the researcher can set the child in a historical context and modify or stabilize the original interpretations, as the case may be. (Steffe et al., 2000, p. 293)

Some authors disaggregate these three phases in four phases, but in essence the main organization is maintained. For example, in the context of educational technology research Reeves (2006, p. 59) considers first the analysis of practical problems by researcher and practitioners in collaboration; second the development of solutions informed by existing design principles and technological innovations; third the iterative cycles of testing and refinement of solutions in practice; fourth the reflection to produce "design principles" and enhance solution implementation.

While educational action research is similar to DBR in terms of its iterative nature, as it may include cycles of identifying a problem, planning, acting, observing, reflecting and reevaluating if needed (Segal, 2009), unlike DBR, it mainly aims at improving educational (e.g., teaching) practices. Allan Feldman (2002, as cited in Segal, 2009) explains that,

> [a]ction research happens when people research their own practice in order to improve it and to come to a better understanding of their practice situations. It is action because they act within the systems that they are trying to improve and understand. It is research because it is systematic, critical inquiry made public. (p. 20)

Even though in my research I played the role of teacher-researcher, I did not precisely aim at changing my own teaching practice, as would have been the case with an action research perspective. My research instead had the aim at identifying design principles that can be used in future research with similar learning goals. DBR additionally aims at contributing with theory elements that might support the development of the field involved in the research. In the following section I focus on details about my own DBR.

II. My Design

To develop my research, I engaged in two cycles of DBR. The first cycle took place in 2017 with four 3rd grade primary school teachers, while the second cycle was conducted in 2018 with three 3rd grade primary school teachers. Both cycles were conducted in Peru. The main goal of developing a first cycle of DBR was to gain some insights for the design of the 2018 design-based research, which I used as the main source of data for my dissertation.

Each research cycle I engaged in consisted of three phases:

- Phase 1: Preparation and Design
- Phase 2: Teaching Experiment
- Phase 3: Retrospective Analysis

Each phase included stages, which varied according to the cycle, as I detail below.

1. Cycle 1

As I mentioned above, cycle 1 was developed in 2017. Here I focus on the description of the phases and stages that were part of the first cycle of DBR.

This cycle included six stages, which occurred in this order:

- Stage 1.1: *Initial Design*
- Stage 1.2: *Exploratory Interview*
- Stage 1.3: *Customized Design*
- Stage 2.1: *Intervention for Teachers*
- Stage 2.2: *Teachers' Teaching and PRE- and POST-teaching Meetings*
- Stage 3.1: *Retrospective Analysis*

Figure 10 summarizes their organization and the way these stages correspond to the phases of DBR.

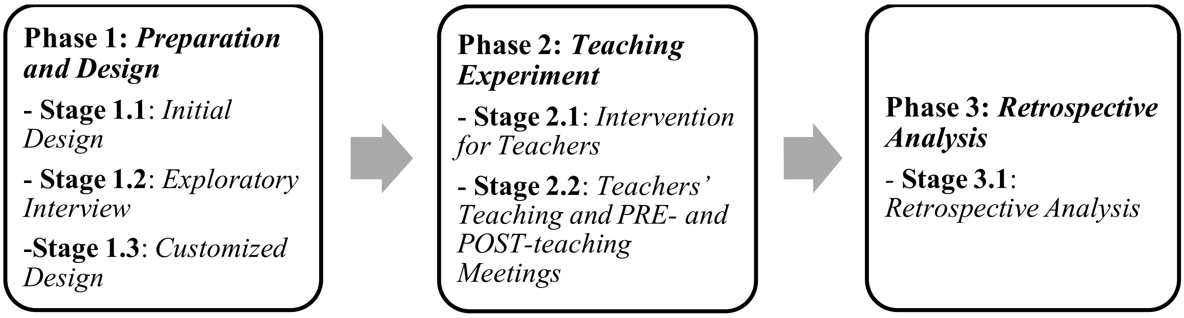


Figure 10. Phases of cycle 1 and the stages embedded in each phase

In the following I focus on each of the phases in this DBR cycle.

1.1. Phase 1 of Cycle 1: Preparation and Design

I divided Phase 1 in three stages. Stage 1.1 was an initial design for the 2017-intervention; Stage 1.2 consisted of an exploratory interview; and Stage 1.3 was a customized design for the 2017-intervention.

Stage 1.1: The initial design

Elementary school teachers in Peru teach most of the school subjects. That means that elementary school teachers' mathematical knowledge is not specialized. Additionally, it is rare that primary school teachers include proof or related topics in their teaching. Hence, I was aware that I needed to design an interventionist design-based research. I began by sketching a first design for the intervention. The main source of reference for my planning was my previous teaching experience and the literature I had reviewed up to that time.

My teaching experience included teaching school students, pre-service teachers in a private university of Peru, my Proof-Based Teaching (PfBT) of division and divisibility to third graders (Vallejo-Vargas & Ordoñez-Montañez, 2014, 2015), the training courses I set up for primary and secondary school in-service teachers as well as those for coaches (teachers who train other teachers). All of the courses I taught included (implicitly or explicitly) proof-related elements, though to a different degree depending on the participants' mathematical background. Part of my prior experience with pre- and in-service teachers and coaches included the teaching of basic elements of logic (e.g., basic set theory; categorical statements; Venn/Euler diagrams), which I usually included before I engaged the pre- and in-service teachers with proving and (in)valid inferences, and my teaching of division and divisibility through PfBT.

The logic part of my teaching was mainly influenced by Irving Copi's work (e.g., Copi, 1973; Copi, Cohen & McMahon, 2014). In my teaching of logic, I did not include truth tables given that my main aim was the teachers' understanding of categorical statements and dis/proving them, which in my view did not require working with truth tables. Another important influence in that subject was Susanna Epp's work (e.g., Epp, 2003, 2009a, 2009b) and notably her remarks on the role of logic when teaching proof (see Chapter 2, Section II).

Proof-Based Teaching (PfBT) is a teaching approach in which mathematics is built in a constructive way. As I pointed out in Chapter 1, it has three important features: (1)

establishing a framework of established knowledge (the toolbox) from which to prove; (2) establishing an expectation that answers and claims should be justified within this framework, which is a *sociomathematical norm* (Yackel & Cobb, 1996) that is set up within the classroom community; (3) proving is the means for constructing new mathematics and it necessarily involves the use of deductive forms of argumentation. These features are described in more detail in Reid and Vallejo-Vargas (2019).

The PfBT approach was directly influenced by my personal mathematical background. When I was a secondary school student I always asked myself where the mathematical formulas I was taught came from. I had the motivation to deduce those formulas for two main reasons: first, I could make sense of the formulas in that way, and second, I did not need to memorize them. In the university, I studied mathematics as my major subject and I learned mathematics by proving theorems. So, my perspective of mathematics was oriented towards seeing mathematics as a constructive theory that involves understanding the origins of theorems.

A structure for the initial design of the 2017-intervention was that it should have two main sections: the first on logic issues and the second on the mathematical content. The design of the first part of the intervention was based on previous mathematics courses I had taught. In concrete, I included a sequence of slides about two main topics: categorical statements and inferences. In relation to the first topic I considered a short review of basic set theory (e.g., inclusion of sets; union and intersection of sets; equal sets; etc.), the logical interpretation of categorical statements, representation of categorical statements with Euler diagrams, and negation of categorical statements (see Chapter 3, Section II.1). The sequence for categorical statements included conditional statements and equivalent statements. The second topic was valid and invalid inferences. My focus for the design of the 2017-intervention was on targeting the teachers' awareness of certain logic-related aspects, for example what the logical interpretation of a universal affirmative statement (UAS) is and what is involved in dis/proving a universal statement (US) (for details, see Chapter 3, Section II). I opted to focus on logic-related issues in the first part of the intervention because I believed that as the teachers learned about categorical statements, their logical interpretation and inferences with them, they would apply those insights to the case of statements framed within the chosen mathematical content in the second part of the intervention.

The mathematical content I focused on in the second part of the 2017-intervention was *Division and Divisibility*[12]. Its design was mainly based on my previous proof-based teaching of this topic, which was taught in terms of the "distribution model" and notably three key notions (fair, whole and maximal distributions, for details see Chapter 1). These key notions made the set of accepted statements (A. J. Stylianides, 2007) that restricted the content of the toolbox from which to start to reason about this content. I chose the mathematical content *division and divisibility* for three reasons: (1) "division" is mathematical content that pupils learn for the first time in primary school and mostly in third grade; (2) I already had experience teaching it in a proof-based way and that facilitated my work given that I had a first design of a proof-based teaching sequence to teach "division and divisibility"; (3) it is a rich mathematical context for justification, in particular, at elementary school level (Ball & Bass, 2003).

I decided to use a proof-based approach to teach this content because I wanted to provide a common proof-based understanding of division and divisibility to the teachers, so that

[12] The topic "Division and divisibility" was only taught in the context of natural numbers during the two cycles of DBR that I implemented.

they could later engage their pupils in similar proof-based understandings (during Stage 2.2.A). I wanted the teachers to have an experience similar to that their pupils would have later. In addition, it allowed me to restructure the teachers' knowledge of the content. I expected them to realize that mathematical knowledge is not a collection of definitions and theorems that appear from nowhere, but that it has a structure. In addition, the proof-based approach to teaching "division and divisibility" has been, in my experience, valuable and meaningful, for both teachers and pupils, given that it supports their development of proof-related skills.

Stage 1.2: Exploratory Interview

The planning of Cycle 1 also included the design of an exploratory interview (Stage 1.2). I engaged the teachers in this interview in order to identify additional initial assumptions that the teachers may have, in addition to those I expected based on my prior teaching experience and my review of the existing literature.

Stage 1.2 had the following goals: learn about the assumptions that the teachers held before the intervention, so that I could modify (in Stage 1.3) the first design of the intervention (made in Stage 1.1), according to these teachers' initial assumptions.

The exploratory interview was made up of two parts: first, an introductory written test; second, a semi-structured interview. The second part was video recorded and all written work was collected. The introductory test included questions about the teachers' mathematical knowledge about the content (division and divisibility), plus some tasks that aimed at investigating some of the teachers' initial assumptions about proving and their understanding of mathematical statements. During the semi-structured interview I wanted to find out whether the teachers used "proof" in their teaching and the criteria they used to evaluate arguments. I asked them questions like: "Have you ever included proof tasks during your teaching?"; "What do you understand by 'proof'?"; "Do you think it is possible that a primary school student proves a mathematical property? (If so, give an example)"; "What role do you think that proof plays at a school level?"

Stage 1.3: The customized design

Based on the findings from Stage 1.2, this stage aimed at modifying the initial design for the 2017-intervention. This included deleting some discussions that seemed not to be crucial for this group of teachers, adding new discussions, and emphasizing some aspects of the first design, for example, by disaggregating some tasks into more elaborated tasks.

1.2. Phase 2 of Cycle 1: Teaching Experiment

The second phase of the second cycle consisted of two stages; namely, Stage 2.1 (the 2017-intervention) and Stage 2.2 (the teachers' teaching and the PRE- and POST-teaching meetings).

Stage 2.1: The 2017-intervention for teachers

Stage 2.1 of the first cycle of DBR consisted of the 2017-intervention for teachers. Four in-service elementary school teachers participated in the study and it was conducted in the private school in Peru where the teachers worked.

As explained in the initial design (Stage 1.1), the intervention consisted of two parts. The first part had a focus on logic-related issues and the second part had a focus on the mathematical content “division and divisibility”. In order to organize the intervention, I used slides and the intervention was as planned.

Stage 2.2: The teachers’ teaching and the PRE- and POST-teaching meetings

Stage 2.2 of the first cycle of DBR was divided into two stages: Stage 2.2.A consisted of the teachers’ teaching in their own classrooms, and Stage 2.2.B was the PRE- and POST-teaching meetings.

After the intervention the teachers taught their pupils the content “division and divisibility” in a proof-based way (Stage 2.2.A). The teachers and I met before each lesson to review the plan for the lesson, which I had designed (PRE-teaching meetings). The teachers then showed their agreement or disagreement with it, and provided suggestions that I included accordingly before sharing a new version with all the teachers. Additionally, after their teaching, the teachers and I met to discuss the difficulties, challenges and positive aspects of their teaching of the lesson (POST-teaching meetings). Based on those meetings, we analyzed whether it was convenient to reconsider the plan we had for the next lesson, whether we should consider other activities, the use of certain terms, etc.

1.3. Phase 3 of Cycle 1: Retrospective Analysis

The third phase of the first cycle of DBR consisted of a retrospective analysis. I watched the videos I collected in the previous phases, reviewed the notes I took during the teachers’ teaching and our meetings, and reflected on those to inform my design of the second cycle of DBR. In contrast to this phase, one change I implemented for Cycle 2 was the addition of Stage 3.2. The main goal was to complement my findings for the effects of the intervention.

Here I describe some of the lessons I learned as a result of my reflections on the phases of Cycle 1 that informed the design of Cycle 2.

Lessons I learned from Phase 1 of Cycle 1

The first phase of Cycle 1 showed me that it was not convenient to conduct the exploratory interview (Stage 1.2) with all the teachers at once. Some teachers were influenced by other teachers’ answers to the questions included in the interview and other teachers did not make efforts to answer as they gave responses like “I think the same”. Hence, in Cycle 2 I conducted the (first) exploratory interview with one teacher at a time.

Lessons I learned from Phase 2 of Cycle 1

In the first section of the 2017-intervention I drew the teachers’ attention to logical principles so that the teachers could use them, but I did not explicitly challenge the teachers’ understandings of these principles. In the 2018-intervention I included challenging the teachers’ initial assumptions as part of the design (see Section 2.1 below). The latter was a more meaningful approach as the teachers could realize and convince themselves of the principles instead of having an external source for it. This was clear during their teaching in Cycle 2. Given that the teachers made sense of the principles, during their teaching the teachers did not only exhibit their personal understandings, but could also guide their students’ realization and sense-making activity.

I also changed the organization of the intervention in my initial design (Stage 1.1). I decided to switch the order of the two main parts. The experience I had with Cycle 1 was that the logic the teachers learned in the first part of the intervention was not actively useful for them when the mathematical content was taught. Besides, there was a lack of emphasis on issues that turned out to be relevant; for example, further discussions about the status of examples when proving or disproving, the relation between conjectures, theorems and proofs, investigating the development of the teachers' understandings of new mathematical reasoning principles, etc. Hence, in Cycle 2 I began with the mathematical content and then focused on logic.

I also created new questions and classroom episodes that I used with the group of teachers who participated in the second cycle, based on the teachers' doubts, hesitations, questions, difficulties, remarks, etc. detected in Cycle 1. For example:

- Discussions 1.4.1 and 1.4.2 of the 2018-intervention were based on my observation of one teacher's teaching in Cycle 1 (Stage 2.2.A), in which the teacher did not seem to notice that a student gave a confirming example when he was supposed to provide a counterexample (see Appendix D1);
- Classroom Episode 9 (see Appendix CE9) was inspired by a discussion the teachers had during the 2017-intervention (stage 2.1) where a teacher accepted a counterexample for the converse as if it were a counterexample for the original statement and I used it in discussion 1.7 of the 2018-intervention to discuss the characterization of counterexamples (see Appendix D1);
- Classroom Episode 13 (see Appendix CE13) was inspired by the teaching of one of the teachers in Cycle 1 (Stage 2.2.A), where the converse problem was manifested and I used it in discussion 2.4 of the 2018-intervention (see Appendix D2);
- Discussion 3 of the 2018-intervention (see Appendix D3) was included based on the inaccurate differentiation that a teacher made between a conjecture and its justification during her teaching in cycle 1 (stage 2.2.A);
- Discussions 4 and 5 (see Appendix D4 and D5) were inspired by an assumption that teachers made in the 2017-intervention that verifying examples are never sufficient mathematical evidence to prove. My decision to include them in the 2018-intervention was also supported by A. J. Stylianides and Ball's (2008) article, where they highlighted the little attention that has been given to the systematic enumeration of all possible cases involved in the construction of proofs, when it is finite.

2. Cycle 2

The second cycle of DBR was conducted in 2018. It included the six stages that were part of Cycle 1, plus a second exploratory interview (Stage 3.2):

- Stage 1.1: *Initial Design*
- Stage 1.2: *First Exploratory Interview*
- Stage 1.3: *Customized Design*
- Stage 2.1: *Intervention for Teachers*
- Stage 2.2: *Teachers' Teaching and PRE- and POST-teaching Meetings*

- Stage 3.1: *Retrospective Analysis*
- Stage 3.2: *Second Exploratory Interview*

Figure 11 shows the phases considered in Cycle 2 and the stages that were embedded in each phase.

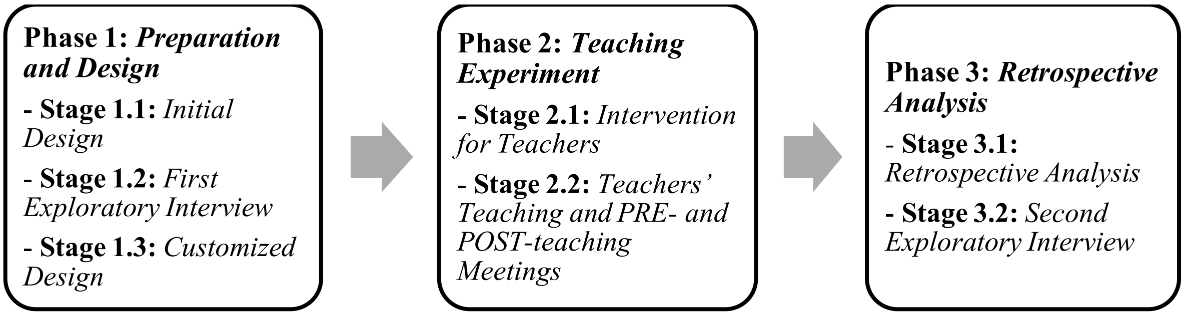


Figure 11. Phases of cycle 2 and the stages embedded in each phase

2.1. Phase 1 of Cycle 2: Preparation and Design

The first phase of Cycle 2 was divided into the same three stages as in Cycle 1: the initial design (Stage 1.1), the first exploratory interview (Stage 1.2), and the customized design (Stage 1.3).

Stage 1.1: The initial design

The preparation and design of the second research cycle was based on three main sources. First, my experience with the first research cycle (see Section 1) influenced my initial design of the 2018-intervention. My decision to focus on the mathematical content "division and divisibility" and the proof-based teaching approach I used to teach the mathematical content are based on the same rationale that I used to establish the design of the 2017-intervention (for details, see Stage 1.1 in Section II.1.1). However, for the 2018-intervention I opted to change the order of the two parts of the intervention and focus on proof-related principles and their understanding instead of only making the teachers aware of logic-related aspects.

Based on the lessons I learned from my implementation of Cycle 1, for the 2018-intervention I switched the order of the two main parts of the 2017-intervention to include first the mathematical content and then the proof-related discussions. My choice to begin the intervention with the mathematical content instead of the proof-related section was mainly based on four reasons: (1) I expected to minimize the influence of the mathematical content in the development of the teachers' principles during the second part of the intervention. By including the first part I targeted the mathematical content which was the frame for the statements considered during the second part of the intervention so it could not hamper the development of the proof-related section. Having focused only on one mathematical content and not others, but also using the same topic to be the main context for the discussions included in the second part had a clear advantage. It diminished the chances that the teachers' attention wandered from our main focus that was the teachers' development of proof-related principles during the second part of the intervention. It means that a lack of understanding of the mathematical content should not have constituted an obstacle for the teachers' performance during those discussions. (2) I wanted to identify, if possible, the teachers' additional initial proof-related assumptions or supporting evidence for those I already identified in Stage 1.2 (see next stage), given that the first part of the intervention included some "simple" proofs that were mainly derivations of the definitions given in this first part. The main purpose was to inform my design of the second part of the intervention. (3) By beginning with the mathematical content section of the intervention, I established a framework from which to justify and begin to reason about proof, which is one of the features of the PfBT approach. The statements used during the second part (for the proof-related discussions) were only framed in this mathematical content. (4) I found support for my decision in the existing related literature. Among the things I read between Cycle 1 and Cycle 2 was Dawkins and Karunakaran's (2016) paper, where they point out that mathematical content can be an influential factor in individuals' proof-related behavior. If the mathematical content plays an important role in people's proof activities, it made sense to first make it as clear as possible so that it mitigated the potential difficulties of carrying out the second part.

I also varied the content of the logic-related part of the intervention to include more proof-related discussions to support the development of the teachers' understanding of mathematical reasoning principles. Reasoning and proving has been recognized to be crucial in the mathematics education of all individuals (e.g., G. J. Stylianides, A. J. Stylianides & Weber, 2017; Mariotti, 2006; National Council of Teachers of Mathematics [NCTM], 2000). In the new intervention I aimed at promoting the teachers' understanding of proof-related principles (e.g., what is involved in dis/proving USs, ESs, and why). Proof is barely (if ever) discussed in professional development programs for elementary school teachers, in particular in Peru. When it is addressed with high school level teachers, it is mostly taught from a formal perspective, so the understanding role of proof (Hanna & Jahnke, 1996; Hanna, 2018) gets lost in the process.

The second main source for my initial design of the 2018-intervention was my review of the literature on proof-related issues, with a focus on difficulties or challenges that individuals face when engaged in proof-related discussions/tasks (see Chapter 2, Section I). For example, it has been widely reported in the literature that individuals of all ages tend to rely on the verification of a few confirming examples when asked to prove an infinite universal statement or when requested to determine the truth value of a generalization (see Chapter 2, Section I.2). Likewise, the problem of the converse has been pointed out as an obstacle for understanding conditional statements and reasoning with them (see Chapter 2, Section I.1). Additionally, there are some challenges when reasoning with "some-statements", given that the existential quantifier "some" is interpreted in different ways, depending the context, for example, either the context is mathematical or it is a real-life one (see Chapter 2, Sections I.6 and I.7).

The third source for the initial design of the 2018-intervention was the proof-related suggestions or considerations made by other researchers (see Chapter 2, Section II). For instance, some researchers have suggested that including tasks with the goal of "translating" statements from one mode of expression to another plays an important role when proving is involved (e.g., Selden & Selden, 1995; Epp, 2009b). I was inspired by this approach when focusing on the logical interpretation of SQ-statements during the 2018-intervention, though my focus was not on the transition from formal to informal modes of representation, or on the other way around. For the 2018-intervention I did not plan to engage the teachers in formal modes of expression because primary school teachers are not familiar with them and never use them in classrooms. Instead, I paid attention to the use of other tools the teachers could use to foster similar understandings; for example, based on my work experience with pre-service teachers, I foresaw the potential in Venn/Euler diagrams since they can be used as a means to understand the logical interpretation of SQ-statements and reason from them. Rephrasing the given statements was another form of manifesting similar understandings. In that respect, I deemed everyday language and "informal" modes of expression to be promising despite of the manifold challenges I was aware that they could add (e.g., Durand-Guerrier, et al., 2012; Epp, 2003; Schleppegrell, 2007, see Chapter 2, Section I). Additionally, G. J. Stylianides and A. J. Stylianides (2009, 2014) designed an instructional sequence with the goal of supporting students' realization of the limitations of empirical arguments when validating mathematical generalizations. They used the cognitive conflict approach to support such a goal. In particular, in the 2018-intervention I included the "monstrous counterexample" they used in their design, which I adapted to my own design. In respect to the role of the context, Epp (2003) has suggested that using examples of statements whose daily-life and mathematical interpretation is the same might support students' development of reasoning skills.

Besides the previous aspects considered in the initial design of the intervention, given that the intervention involved the use of the PfBT approach, its design needed to include the features of PfBT. As for the case of the framework of established knowledge, the need to provide justifications for answers and claims was an essential sociomathematical norm (see Yackel & Cobb, 1996) that I introduced from the intervention's outset and kept while the teachers shaped their understanding of "proof". Within the design of my intervention, this norm was expected to support the teachers' overcoming of the cognitive conflicts I planned to provoke in their understanding process. With this I aimed to trigger the teachers' need to find mathematical reasons for their new proof-related assumptions.

There were some additional concrete changes I considered for the design of the 2018-intervention as a result of my reflections on Cycle 1 of the DBR. In general, some of the modifications were given in terms of the content as well as my teaching approach for both parts of the intervention.

The design of the first part of the 2018-intervention

The focus of the first part of the 2018-intervention was on the mathematical content *Division and Divisibility*. This mathematical content was taught in a proof-based way. This part of the intervention targeted the teachers' understanding of several basic properties and theorems of division and divisibility. For instance, understanding why:

- the remainder must be smaller than the divisor;
- the maximal remainder in a division with divisor "n" is equal to "$n-1$";
- zero is divisible by any non-zero natural number;
- any natural number is divisible by 1;
- any non-zero natural number is divisible by itself;
- all the numbers divisible by 6 are divisible by 3; etc.

In Cycle 2 the teachers proved more divisibility properties than those in Cycle 1, by following the proof-based approach.

The first part of the 2018-intervention included the analysis of actual classroom episodes, most of which were taken from the data collected in Cycle 1. They involved the analysis of real (pupils and teachers') answers to tasks that were similar to those that the teachers in the 2017-intervention solved. The goal of these analyses was to familiarize the teachers with real arguments, which would provide a first approach to criteria for how to decide whether an argument proved a mathematical statement about division and divisibility, without getting into details that were expected to be filled during the second part of the intervention. Most of the arguments presented in this part of the intervention were direct arguments.

The design of the second part of the 2018-intervention

Both parts of the intervention included actual classroom episodes for analysis, which were adapted from the data I collected in Cycle 1. In addition, I decided to divide the second part of the intervention into discussions, which involved more engagement of the teachers. This part of the 2018-intervention consisted of nine discussions, each with a particular focus (see Table 3). Every discussion (except for Discussions 4, 5, 8 and 9) was divided in sub-discussions that aimed specific related goals. For example, Discussion 1 had a focus on false universal affirmative statements and was split in ten sub-discussions. Each sub-discussion involved the use of one slide. The first sub-discussion (Discussion 1.0) aimed at exploring the teachers' initial approaches to disprove a universal affirmative statement (UAS); the second sub-discussion (Discussion 1.1) explored whether the teachers initially accepted repetitive arguments as a valid disproof; the third sub-discussion (Discussion 1.2) had a focus on whether two or more students needed to provide exactly the same counterexample to disprove a UAS; and other related topics (for details[13], see Appendix D1). Not all discussions have the same number of sub-discussions. For example, discussions 4, 5 and 8 consisted of only one main discussion.

Table 3. Discussions during the 2018-intervention and their main focus.

Discussion #	**Main Focus**
Discussion 1	False universal statements: valid disproofs; counterexamples; characterization/description for counterexamples.
Discussion 2	Universal Conditional Statements: their connection with universal statements; the problem of the converse.
Discussion 3	Conjectures, Justifications and Mathematical Truths: their connection; false conjectures with confirming examples; a false conjecture with an extreme counterexample; an open conjecture.
Discussion 4	Universal statements and the number of cases involved.
Discussion 5	Universal statements that involve the set of natural numbers in their formulation and the number of cases involved.
Discussion 6	True universal affirmative statements: valid proofs according to the number of cases involved.
Discussion 7	Existential statements: negations of universal statements; equivalent statements; (non) allowed inferences; negation of existential statements. NO-statements: equivalent statements; negation of NO-statements.
Discussion 8	A conditional statement that is not equivalent to a universal statement.
Discussion 9	A review

Unlike in Cycle 1, the second part of the 2018-intervention was planned to focus on overarching features: the logical interpretation of single-quantified (SQ-) statements, the development of mathematical forms of reasoning and their understanding, and dis/proving of SQ-statements. Now I proceed to explain what I considered in each case.

[13] Appendix D1 contains the tasks considered in each of the sub-discussions included in discussion 1

The logical interpretation of SQ-statements

For each type of SQ-statement the proof-related discussions first aimed at understanding its logical interpretation (see Chapter 3, Section II.1). Without a clear understanding of the logical interpretation of a SQ-statement, it may be hard to consciously grasp what kind of mathematical evidence it takes to dis/prove it. The teachers' understanding of the logical interpretation of a SQ-statement included being able to answer the following questions: "Do you understand what the statement is about?", "Are you aware of what elements the statement refers to and what it is claimed about those elements?", "What is the set of analysis?", "Is the statement about *all* or *some* of the elements in the set of analysis?", "Can you rephrase the statement given in different equivalent ways?". Grasping the logical interpretation of a SQ-statement was the first step towards engaging the teachers in further proof-related discussions about those statements.

The development of mathematical forms of reasoning and their understanding

Every discussion included during the 2018-intervention had a specific goal in relation to the development of mathematical forms of reasoning. For instance, Discussion 1 targeted the understanding of why one counterexample was sufficient to disprove a universal statement (US). Additionally, it aimed at developing a general characterization for counterexamples. Discussion 2 focused on the understanding of the connections between universal conditional statements (UCSs) and USs, as well as understanding that once a true US is given, nothing can be conclusively inferred about the truth value of its converse. In Discussion 3 the topic was what it takes for a conjecture to become a mathematical truth; that the verification of some (but not all) of the cases involved in an infinite US did not suffice to prove it, unless the confirming examples were part of a generic proof. Discussion 4 drew attention to the number of cases involved in the set of analysis and understanding proof by exhaustion. Discussion 5 targeted the understanding that USs that involved the set of natural numbers in its formulation did not necessarily imply that it involved an infinite number of cases. Discussion 6 had a focus on understanding what constituted a mathematical justification (a proof) according to the number of cases involved in a universal statement. Discussion 7 aimed at promoting the understanding that given a true existential affirmative statement (EAS), it does not necessarily imply that its respective existential negative statement (ENS) is true; the negation of a US is an existential statement (ES) and not a "no-statement" (NO-S); the sufficient mathematical evidence to prove an ES is a confirming example. Discussion 8 had as a goal to draw attention to cases of conditional statements (CSs) that may not be equivalent to USs. Understanding these forms of reasoning entailed being able to provide mathematical reasons for why they held.

Dis/proving of SQ-statements

In the process of developing mathematical forms of reasoning, the initial design of the 2018-intervention included different opportunities to dis/prove SQ-statements, which was supported by the emerging forms of reasoning.

In particular, the second section of the intervention had as a main goal to engage the teachers in proof-related discussions that could be relevant for them to lead similar in-the-moment mathematical discussions in their own classrooms (Stage 2.2.A). Some of the aspects considered was to enable the teachers to assess their students' spontaneous arguments, guide their students' work without providing direct answers, establish and activate specialized mathematical language in class, etc.

An important factor considered in order to achieve these goals was to provide opportunities for the teachers to anticipate students' answers (which included the case of arguments) by having access to and analyzing different students' answers, for example those collected in Cycle 1 and in my own previous experiences with proof-based teaching. Some of these answers took the form of classroom episodes, others were presented on slides used for the intervention and some others became activities.

The teachers engaged in proving and disproving existential and universal statements; however, the main focus of proving universal statements with infinite cases involved was on direct proving.

Stage 1.2: The First Exploratory Interview

The planning of my second research cycle included a first exploratory interview (Stage 1.2), just like for the case of Cycle 1 (see Stage 1.2 in Section II.1.1 above). Like in Cycle 1, Stage 1.2 of Cycle 2 also aimed at identifying some of the teachers' initial assumptions.

In contrast to Cycle 1, the order of the sections in this stage were shifted: the first part was the semi-structured interview. Unlike the first part of Stage 1.2 in Cycle 1, this stage in Cycle 2 also had a focus on identifying the teachers' assumptions related to mathematics, its teaching and learning. The second part consisted of a task-based interview (Maher & Sigley, 2014). This included following-up questions that I asked with the purpose of checking that the teachers understood the tasks included in the test and, furthermore, identifying the teachers' additional initial assumptions. Unlike Cycle 1, I conducted both parts of Stage 1.2 with one teacher at a time.

Both parts were video-recorded and the first part was transcribed. For the second part, a list of tasks was handed out to the teachers with the purpose of sketching written answers. Whenever I noticed hesitation in the teachers' answers, I interrupted the interview to ask further questions that could provide any insight about the origins of such hesitation (e.g., "Do you have doubts about that question?", "What part of the task you do not understand?", "Are you hesitating about your answer?", "Do you need me to explain something about the task?"). When the teachers' answers were too short or not clear enough, I asked them back to further clarify (e.g., "Could you please expand in your answer?", "Could you please give details about it?", "What do you mean exactly by 'that'?")

For my analyses I considered both the teachers' written and oral responses during the introductory test. A. J. Stylianides (2019) found out that the oral mode of representation was more likely to meet the standard of proof than the written mode. He highlighted that "by considering only one mode of representation and ignoring the other, each study individually would have reported an incomplete picture of students' constructed proofs, for apparently it matters whether students present their perceived proofs orally or in writing" (p. 22). The mode of representation used in arguments has been underscored by G. J. Stylianides and A. J. Stylianides (2020) as a factor that influences the students' proof construction process.

Stage 1.3: The customized design

In this stage, and based on my observations during Stage 1.2, I made changes to my initial design, according to the teachers' needs. As a result, I put more emphasis on the discussion I had initially designed for the understanding of existential statements, and specifically for the understanding that a statement of the form *"Some X are Y"* does not necessarily imply that *"Some X are not Y"*. In addition, given that one of the teachers had expressly mentioned she used the term "counterexample", I included two additional overarching features: the development of key proof-related concepts and the use of different content in statements.

The development of proof-related concepts progressively

The 2018-intervention had also a focus on developing some key proof-related concepts. There were three key terms whose concepts were expected to (at least) begin to be shaped during the intervention: "counterexample", "dis/proof", and "some". In order to do this, factors such as the teachers' initial common-language (or possibly other sources) interpretations of those terms needed to be considered. In general, the teachers were given initial inputs, then some other approaches were taken to shape the initial input. For example, in order to develop the existential quantifier "some" concept, a cognitive conflict approach was taken. The concept "dis/proof" was developed gradually throughout the 2018-intervention, given that it involved understanding the dis/proving of different SQ-statements (see Chapter 3, Section II). The concept "counterexample" was targeted through the use of an initial incomplete input for it, and irrelevant and confirming examples that played the role of shaping the initial input.

The use of different content in statements

The content of a problem may influence the display of logical skills (Hawkins et al., 1984). I call an "abstract statement" a statement that is formulated in general terms, hence its content is irrelevant and its truth value is impossible to determine. The statements *"There are X that are not Y"* and *"No X is Y"* are examples of abstract statements since *X* and *Y* might be anything as nothing is specified about sets *X* and *Y*. I call a "familiar statement" a statement whose content is close to the participants' background knowledge. The statements *"If it is a person, then it is a mortal"* and *"Some cats are independent animals"*, which are framed in a real-life context, are examples of familiar statements.

Hawkins et al. call syllogism problems "in which premises described mythical creatures foreign to practical knowledge" (p. 586) *fantasy problems*. They give the following example:

> Every banga is purple.
>
> Purple animals always sneeze at people.
>
> Do bangas sneeze at people? (p. 587)

Hawkins et al. (1984) show that preschool children are able to reason about fantasy problems deductively, as these problems eliminate the possibility of being influenced by practical knowledge.

Inspired by Hawkins et al., I introduced what I call *imaginary statements*. These are statements in an imaginary context, like *"Each «bogui» is a «fantaslopitocus»"*. Imaginary statements are not completely *abstract* and neither are they concrete like mathematical or *familiar* statements. Imaginary statements are in-between concrete and abstract statements. Furthermore, there might be cases of "partially-imaginary" statements, such as *"If a number is odd, it is «poponupulus»"*. In those statements only one of the sets involved is known. We know what odd numbers are, though what "poponupulus" numbers are remains unknown.

Like in the case of statements that are part of fantasy problems, the truth value of imaginary statements is indeterminable. Given that at least one of the two sets involved in the statement is unknown, its truth value cannot be determined. In contrast to fantasy statements, imaginary statements do not refer to mythical creatures that might still be attached to practical knowledge and rejected by some individuals. For example, some individuals might claim that "Purple animals always sneeze at people" is false and refuse to reason with that statement because Barney does not always sneeze at people. At most, imaginary statements like *"Each «bogui» is a «fantaslopitocus»"* do not specify what a "bogui" and a "fantaslopitocus" are (e.g., it does not state whether they are animals).

Similar to the case of fantasy statements, the no interference with practical knowledge can make that imaginary statements draw attention to the logical structure of the statement and to the general aspects of the task in which it is involved.

The First Exploratory Interview included a task[14] in which an imaginary statement was given as well as a list of twelve statements from which to choose those that stated the same thing as the given imaginary statement. Given that the teachers' solutions revealed their current assumptions about single-quantified statements, I decided to include this as a feature that the 2018-intervention should have.

2.2. Phase 2 of Cycle 2: Teaching Experiment

The second phase of the second cycle consisted of two stages: Stage 2.1 (the 2018-intervention) and Stage 2.2 (the teachers' teaching and the PRE- and POST-teaching meetings).

The teachers and the school context

The study I report here was conducted in a private school in Peru that belongs to the same chain of private schools as the school where Cycle 1 took place. The main language used in those schools is Spanish[15]. Three in-service elementary school teachers participated in the study: Andrea, Gessenia and Lizbeth. By midyear of 2018, the teachers had 8, 10 and 15 years of teaching experience, respectively. In this private school only four teachers taught a third-grade class: Andrea, Gessenia, Lizbeth, plus another teacher who could not participate in the study. Every third-grade class had four mathematics sessions per week. There were about 27 pupils per third-grade classroom. Most of Andrea's teaching experience was in second grade; Gessenia's was in second and third grades; and Lizbeth's was in first grade. They all shared a common initial perspective about mathematics: mathematics is mainly about problem solving, for which most likely they did not mean solving complex non-standard problems, but procedural problems and doing exercises. That, presumably, was a direct influence from the philosophy of mathematics that mathematics teachers followed in the school.

I chose a private school over a public school because they offered me to facilitate my work with the teachers and pupils. The school authorities were flexible with the time I

[14] For details, see Task 2 of Section II in Appendix EI1-P2.

[15] Some schools in Peru teach all or some subjects in other languages, such as English.

spent with the teachers and allowed me to be present during the teachers' teaching and video-record those sessions. They also provided rooms for our meetings and materials like printouts for the pupils. The teachers were also highly engaged with the work we developed.

The three teachers agreed to participate in the intervention. They accepted that I use their real names. They also agreed on being video and audio recorded during the development of my work with them.

Stage 2.1: The 2018-intervention for teachers

The Cycle 2 intervention was conducted in 2018. It consisted of two parts, the first part focused on the mathematical content and the second on proof-related discussions.

The two parts of the intervention together lasted approximately 22 hours (9 hours 30 minutes and 12 hours 45 minutes, respectively). The intervention was video-recorded and among the materials I used slides that were the main resource to organize my work with the teachers, as well as activities and classroom episodes. The slides were printed out and handed out to the teachers at the end of every session, as were the activities and classroom episodes when needed. Only the teachers' solutions to the activities were collected.

After each session of the intervention I reflected on the difficulties/challenges/obstacles the teachers experienced during the session and on that basis I decided to keep or modify the design of the intervention for the next session.

I designed Stages 1.2 and 2.1 to shed some light on the teachers' initial proof-related assumptions and given that particularly Stage 2.1 engaged the teachers in the 2018-intervention, both stages supplied the data I needed to focus on my Research Question 1 (What are the effects of an intervention focused on proof-related issues and PfBT on in-service primary school teachers' ability to engage in proving and their understanding of the nature of proving and its role in Mathematics and Mathematics learning?) and the related Question 1a (Why do these effects occur? What features of the intervention led to these observed effects? Why did the intervention have these effects?).

Stage 2.2: The teachers' teaching and the PRE- and POST-teaching meetings

Stage 2.2 consisted of two sub-stages: the first one (Stage 2.2.A) had a focus on the teachers' teaching and the second one (Stage 2.2.B) on the PRE- and POST- teaching meetings. The teachers' teaching took place in the mornings, during the pupils' regular school timetable; while the PRE- and POST- teaching meetings took place in the afternoons, once all the teachers had finished their daily teaching. Each teacher taught four mathematics sessions per week, and we all met five afternoons each week. In total and in the context of the 2018-intervention, every teacher taught 13 lessons in her own classroom and each lesson lasted around 80 minutes. This was the regular time the teachers used for their math lessons. On the other hand, the PRE- and POST- teaching meetings were 17 in total and each lasted 60 minutes, 3 days a week, and 90 minutes, 2 days a week. The time assigned to the lessons and the meetings was allotted by the school.

The role I played during Stage 2.2.A was mainly to accompany the teachers during their teaching, observe and take notes; however, there were a few occasions (one or two of the first lessons) where I intervened in each teacher's classroom. The main aim of my interventions in the teachers' classrooms was to give confidence to the teachers, and show them that it was possible that their students could get engaged in reasoning and proving while learning about division and divisibility in a proof-based way. One motive to engage myself in this experience was the skepticism that the teachers explicitly expressed during our first PRE-teaching meeting, at the end of Stage 2.1. They showed their concern about the unlikelihood they anticipated for their students to come up with mathematical properties and their proofs. Hence, it was a good opportunity for me to contribute in each teacher's classroom, but also for the teachers to see me as a supporting participant and not as an evaluator, which I believe they were more afraid of, at least at the beginning of Stage 2.2. This experience clearly contrasted with my participation in the first cycle, when I did not actively participate in the teachers' classrooms.

The advantage of following a pre-determined proof-based sequence of tasks for their teaching was that freed the teachers from making decisions about an appropriate order in which to organize the tasks while they were teaching. This was possible since I counted on a PfBT theory for the content "division and divisibility", which the teachers had learned about during Stage 2.1 and it served as a framework for them to engage their own students in proving.

Due to the overlap of some lessons, I did not personally observe all of the teachers' lessons; however, I video-recorded and transcribed all teachers' teaching sessions. This, as well as my notes and the photos I took of the teachers' annotations on the whiteboard, were the main source of data for my analysis of this stage.

After the teachers' teaching I watched the videos of the lessons where I could not be present whenever possible. While watching those videos my attention was focused on detecting difficulties the teachers could have experienced while teaching. That information, together with my observations of lessons where I could be present, I used to prepare tasks that took the form of classroom episodes or quick questions, where I addressed, for instance, language imprecisions such as "if I am asked to complete the statement 15 is divisible by …, I am asked to give the numbers divisible by 15", "a number is divisible by another when I can divide the first by the second", "a counterexample proves that a statement is false". This invited the teachers to reflect on their practice and important proof-related activities.

The goals of the PRE- and POST- teaching meetings were the planning of the lessons and the reflection on what was done during their teaching, respectively. Since I already counted on a well-developed proof-based teaching sequence for division and divisibility, the planning of the PRE-teaching meetings consisted of suggestions for changes that the teachers could make based on a draft proposal I submitted to them in every meeting. The POST-teaching meetings focused on the teachers' reflection on what they experienced in their own classrooms. For example, they were expected to share with the other teachers whether they believed the students could follow the lesson design easily or not, as well as suggest new changes for next lessons based on their recent observations. Additionally, they were encouraged to point out their own strengths and weaknesses during their teaching. Besides the focus set for the PRE- and POST- teaching meetings, whenever it was possible, we continued with the teachers' reflection on and preparation for proof-related issues that could not be completed during Stage 2.1 and could still support the teachers' teaching.

Stage 2.2 provided me the data I needed to answer my Research Question 2 (How are the effects of such an intervention reflected in the teaching of in-service primary school teachers while teaching in schools?).

2.3. Phase 3 of Cycle 2: Retrospective Analysis

The third phase of the second cycle of my DBR consisted of systematically going through all the data I had collected. I divided Phase 3 into two stages: retrospective analysis (Stage 3.1) and the second exploratory interview (Stage 3.2).

Stage 3.1: Retrospective Analysis

This stage involved watching all the videos from the previous two phases, transcribing them, finding connections between the teachers' assumptions in hindsight, carefully paying attention to the teachers' utterances, their gestures, their choices (e.g., during their teaching), reflecting on the previous aspects by going forwards and backwards to make sense of the development of their assumptions.

I organized stage 3.1 into six steps, with these goals:

1. identify and organize the teachers' proof-related assumptions;
2. determine the teachers' *initial* proof-related assumptions;
3. identify any deviation from the initial assumptions;
4. track any evidence that supports or contradicts the findings;
5. identify possible factors that might have triggered the changes;
6. revisit the data until no further assumptions, deviations and factors emerge.

Next, I include a description of the steps I took to perform the analysis of my data.

Step 1: Identify and organize the teachers' proof-related assumptions

Before engaging in determining the teachers' *initial* proof-related assumptions, I tracked all the teachers' proof-related assumptions from Stages 1.2, 2.1 and 2.2. In order to do so, I first tried to isolate every teacher's proof-related assumptions. This was occasionally tricky because during the intervention the teachers talked at the same time or some did not share much. For this reason, it was important to include a stage where the teachers could express their personal assumptions at a more individual level (see Stage 1.2 in Figure 11 above).

I organized the teachers' related assumptions in a timeline in the order they emerged so that later I could make a distinction between initial and modified assumptions.

It is important to notice that each teacher had more than one "assumption development timeline" given that each had a specific focus of attention; for example, there was one "assumption development timeline" for proving USs, another for negating ESs, other one for understanding the logical interpretation of ESs, etc.

Step 2: Determine the teachers' initial proof-related assumptions

Given that my interest was on the effects of the intervention, it was crucial to find out what were the teachers' *initial* assumptions. In that aspect, the research literature played an important role since it gave me directions from where to start when I looked into the teachers' "assumption development timeline". For example, the literature in respect to the relationship between a UCS and its converse shows that most individuals tend to presume that a UCS implies its converse, and vice versa (see Chapter 2, Section I.1). Likewise, it is very common that individuals provide an empirical argument when requested to prove an infinite US and some perceive them as sufficient evidence (see Chapter 2, Section I.2).

A relevant observation here is that the order in which the assumptions were exhibited in the “assumption development timeline” did not necessarily imply that the first-to-show assumptions were the teachers’ *initial* assumptions. In some cases, the process of determining the teachers’ initial assumptions became difficult because some of them were more fundamental than others and in some cases, they were accessible only after the intervention for teachers (i.e., in Stage 2.2). For instance, that was the case for the understanding of “no-statements” (see Chapter 6, Section II.2.4). In addition, initial assumptions might be explicit or implicit assumptions. In that sense, it was crucial to pay close attention to the teachers’ utterances and interpret carefully what they said.

Step 3: Identify any deviation from the initial assumptions

Besides looking for clear (explicit) signs of changes, this also involved paying attention to any of the teachers’ manifestations of doubt or hesitation. For example, if a teacher initially assumed that *“If X, then Y”* implied that sets *X* and *Y* were equal, a manifestation of deviation of this initial assumption included that s/he posed a question like “Is it possible that *X* and *Y* are not equal?”, or something subtler like “Can we have an element in *Y* that is not an element in *X*?” (see e.g., the case of Gessenia in Chapter 5, Section I.1.1.2).

Step 4: Track any evidence that supports or contradicts findings

Collecting evidence that might support the teachers’ initial assumptions and deviations from those assumptions was also important. It was not sufficient to point to the teachers’ assumptions, but it was also relevant to track the evidence that supported those findings, and evidence that brought them into question and required reconsideration. This evidence included utterances that the teachers shared in a discussion, their solutions to tasks, changes in their behavior (e.g., unusual silence when a teacher is normally participative during discussions, gestures that reveal confusion), and questions they asked. All these were used when determining whether the intervention had an impact on the teachers’ initial assumptions.

Step 5: Find out possible factors that might have triggered the changes

In order to identify potential factors for deviations from initial assumptions, I needed to return to my data and focus on the teachers’ justifications for their “new” assumptions. Were those explanations tied to the intervention? How?

With this step I expected to investigate the nature of the changes that were manifested during the intervention because they were a result of the intervention, or changes that were superficial, perhaps they were utterances that the teachers repeated from other teachers.

In order to engage in this step, it was important to make possible connections between related assumptions. For example, assumptions related to existential statements may be linked to assumptions related to the negation of universal statements; assumptions about the falsity of existential statements may be related to assumptions related to the truth of such statements; etc. Hence, it was not only important to look into the teachers' "assumption development timeline", but also across them.

Step 6: Go back and forth until the data is exhausted

In order to perform the previous steps in detail, I needed to go back and forth my data several times, share my observations with a third party (in this case my supervisor), with the goal to eliminate potential bias in my analysis.

This procedure can reveal that the teachers might hold interconnected assumptions that may not be clear enough to be seen at first glance. For example, the way teachers interpret the quantifier "some" might affect their negation of "all-statements" and also the way they refute "some-statements" (see Chapter 5, Sections I.4 and II.2).

In order to complement the findings from Stage 3.1, I engaged the teachers in a second exploratory interview, which constituted Stage 3.2.

Stage 3.2: The second exploratory interview

This interview with the teachers was planned with the goal of collecting the teachers' impressions on what has been done along the previous stages. It consisted of three parts: first, we (the three teachers and I) engaged in an open conversation where the teachers shared some of the general insights they gained during their participation in my study; second, I gave the teachers back the written test they completed in Stage 1.2 so that they could reflect on their initial answers and change them if they noticed that there was something wrong or inaccurate in them; third, I sent the teachers a list of questions that they could answer individually at home. The questions were similar (but not the same) to those I used in the first part of the interview. The main difference is that in this part they would have time alone to reflect on their answers and as it was a "private" task, they could feel more confident to freely express their ideas. Among those questions I included:

- (Question 1) Do you notice any changes in the way you see Mathematics after the training course[16] period? If your answer is "yes", describe how this changed as a consequence of the training course.
- (Question 2) Have you learned something new in each stage? Describe exactly what.
- (Question 3) What specific moments of the training course were more meaningful to you? Explain why. It is possible that our meetings were not meaningful to you. Point that out and explain why.
- (Question 4) Write down three characteristics from experiences you had along the training course that directly affected your teaching practice (for example, strategies you didn't use and now you do as you see positive effects in your students; knowledge you didn't have and that you realize that is useful in your classes; etc.). Include examples and details from the experiences you consider relevant.
- (Question 5) Now that you have seen an example of what it means to teach in a proof-based way, a) Do you think you could have taught in this way without

[16] With the teachers I used the expression "training course" as it was presented as such by the school authorities.

having before the first stage[17] of our training course? What about the second stage?; b) Do you think it is important that your students keep having new experiences with this way of learning mathematics? Why?

All three previous parts of Stage 3.2 facilitated me a direct access to the effects of the intervention in the teachers' ability to engage in proving and their understanding of the nature of proving and its role in mathematics and mathematics learning (my Research Question 1) from their own perspective.

Both Stages 3.1 and 3.2 helped me to reflect on design principles I would suggest for future interventions with similar goals, which is directly related to my Research Question 3 (What design principles for a teacher development intervention focused on Proof-Based Teaching [PfBT] in primary schools can be abstracted from two cycles of such interventions?), as well as suggestions for future research.

[17] Here "the first stage" means the 2018-intervention (i.e., Stage 2.1 in my design), whereas "the second stage" means the PRE- and POST-teaching meetings (i.e., Stage 2.2.B in my design).

Chapter 5: Findings and Interpretations from Cycle 2

Chapter 5 includes the findings from Cycle 2 of my design-based research (see Chapter 4). Hence, whenever I refer to "the intervention" in this chapter I mean the 2018-intervention. These findings will help me answer my research questions, which are the following:

- RQ1: How do in-service primary school teachers' assumptions related to dis/proving change while engaged in an intervention focused on Proof-Based Teaching (PfBT) and understanding the nature of proving?
- RQ1a: What features of the intervention led to the observed changes?
- RQ2: How are the in-service primary school teachers' assumptions that changed visible during their teaching in schools?
- RQ3: What design principles for a teacher development intervention focused on Proof-based Teaching (PfBT) in primary schools can be abstracted from two cycles of such intervention?

The findings focus on the teachers' assumptions, the way the assumptions developed during the intervention and aspects of the intervention that possibly contributed to that development. As I present the findings I will also discuss connections between my observations and prior research.

The development of the teachers' assumptions was a long process. Specific assumptions mark steps in the development, and episodes that prior to, during and after the intervention provide evidence of the teachers' assumptions at each time. The teachers' initial assumptions were mostly (but not only) observable before the intervention. During (and sometimes after) the intervention the teachers revealed the development of their assumptions, either as changes of their assumptions or changes of the rationale behind their assumptions. Some of the teachers' new assumptions were observable in particular after the intervention, during their teaching. Notably, the teachers' teaching exhibited the nature of the teachers' current proof-related understandings.

My reporting of the findings is thematic, not strictly chronological. I divide this chapter in three main sections, according to the type of statements that were the main focus of our discussions: the teachers' assumptions about universal affirmative statements (Section I), the teachers' assumptions about existential statements (Section II) and the teachers' assumptions about universal negative statements (Section III).

Each section is divided in sub-sections that are more specific topics within the section (e.g., confirming examples and their status when proving universal affirmative statements). For every sub-section, I focus on the assumptions of each teachers separately. Not all teachers had observable assumptions related to each specific topic. This means that a sub-section could be divided into one, two or three additional embedded teacher sections.

I begin each teacher's section with a table that summarizes her assumptions. The table includes: the teacher's assumptions about the specific topic, whether it is an initial or new assumption, and when the assumption was first observed (before, during or after the intervention).

At the end of each section I include a summary and a figure that shows not only the development of the teacher's assumptions, but also the elements or aspects of the intervention that might have supported or triggered changes.

I. The teachers' assumptions about Universal Affirmative Statements

In this section I consider the development of the teachers' assumptions about universal affirmative statements (UASs[18]) and particularly those related to four main themes: (1) Logical Interpretation of Universal Affirmative Statements; (2) Disproving of Universal Affirmative Statements; (3) Proving of Universal Affirmative Statements; and (4) Negation of Universal Affirmative Statements.

During the 2018-intervention for teachers, different aspects of UASs were discussed throughout the nine discussions that were part of it. The falsity and disproving of UASs was the main focus of Discussion 1. Universal Conditional Statements (UCSs) were the dominant topic of Discussion 2, with more emphasis on false UCSs and therefore both Discussions 1 and 2 were closely related. True and false universal statements (USs) and the way their *epistemic value*[19] (see Duval, 2007) may change, for example from a conjecture to a mathematical truth by way of a mathematical proof, was the focus of Discussion 3. The number of cases involved in UASs was the focus of Discussions 4, 5 and 6, with special attention to their proof in Discussion 6. The negation of USs was debated in Discussion 7; however, the main goal of such discussion was to introduce existential statements, which is the focus of Section II. Noticing that not all conditional statements are universal statements was the focus of Discussion 8. A general review of the previous discussions was the main aim of Discussion 9.

In general, the form of the statements may influence the way they are interpreted, as I show in this chapter[20]. In particular, Universal Affirmative Statements can take different forms[21], but the intervention had a focus mainly on three: *"The Xs are Ys"*, *"All X are Y"*, *"If X, then Y"*.

This section is divided into four sub-sections, addressing the development of the teachers' assumptions related to the four themes listed above.

1. Logical Interpretation of Universal Affirmative Statements

In this section I focus on a single theme directly linked to the logical interpretation of universal affirmative statements[22]: the relation between a UAS and its converse.

[18] For more details on what I mean by *universal affirmative statements*, see Chapter 3, Section II.

[19] Duval (2007) explains that the *epistemic value* of a proposition/statement "is closely connected to way somebody understands the content of a proposition: it depends on the subject's knowledge basis. For example, this way of understanding can be 'theoretical,' that is with a background of definitions, theorems and deductive practice, if the subject is an expert mathematician, or it can be only 'semantical,' that is reflecting ordinary language understanding, if the subject is a young learner. For example, any proposition whose content focuses on mathematical properties which can be immediately seen on a figure (parallelism, perpendicularity, etc.) can have quite different epistemic values: visually obvious for the student but only possible or, maybe, impossible from a mathematical point of view." (p. 138)

[20] In concrete, I observed this for the case of the existential "some-" and "there-exist-" statements and the universal negative "for-every-" and "all-" statements as I show in Sections II and III, respectively. Hence, it is possible that by considering other UASs, other assumptions might have emerged.

[21] For other forms of UASs, see Chapter 3, Section II.2.

[22] For details on what I mean by *logical interpretation of universal affirmative statements*, see Chapter 3, Section II.1.

Specifically, I concentrate on the development of Andrea's and Gessenia's assumptions related to this topic, their differing initial assumptions and the aspects of the intervention that may have influenced their development.

1.1. Relation between a UAS and its converse

From a mathematical perspective, a UAS and its converse do not state the same thing and hence a true UAS does not necessarily imply that its converse is true or false. Andrea and Gessenia began the intervention with different assumptions about UASs and their converses and how these statements were related to each other. On the one hand, Gessenia assumed that a UAS and its converse state the same thing, which she made sense of by assuming that a UAS implied its converse, and vice versa. On the other hand, Andrea assumed that if a UAS was thought to be true, it followed that its converse was false.

Section 1.1.1 is about Andrea and the way her assumptions evolved throughout the intervention. I begin by pointing to Andrea's initial related assumptions, then I show that Andrea used some of the concepts introduced during the intervention to justify her initial assumption that a UAS and its converse did not state the same. Finally, I focus on the change of Andrea's initial assumption that by assuming a UAS to be true, it follows that its converse is false. I explain the role that the introduction to true biconditional statements during the intervention seems to have played in her change of assumption.

Section 1.1.2 is about Gessenia and her resistant-to-change assumption that a UAS and its converse stated the same. I show how Gessenia's assumptions gradually evolved. I explain the way familiar-context examples of true UASs with false converses constituted conflicting cases and played an important role in Gessenia's reconsideration of her initial assumption during the intervention.

1.1.1. Andrea's assumptions

This section has a focus on two understandings that Andrea developed during the intervention. The first is related to the rationale Andrea used to explain why a UAS and its converse do not state the same. The second understanding is centered on Andrea's reconsideration of her initial assumption that a true UAS always implies a false converse.

Table 4 (below) includes Andrea's assumptions about the relation between a UAS and its converse (first column) and whether it is an initial or new assumption as well as when it was first observed (second column).

Table 4. Andrea's initial and new assumptions about the relation between a UAS and its converse

Assumption[23]	**Type of assumption (when first observed)**
bAA1: *"All X are Y"* can be represented as set X included in set Y, with $Y - X \neq \emptyset$	Initial assumption (before the intervention)
bAA2: If a UAS is (assumed to be) true, its converse is false	Initial assumption (before the intervention)
bAA3: A UAS and its converse do not state the same	Initial assumption (before the intervention)
dAA2[1]: *"All X are Y"* can be represented as X included in Y and $Y - X$ may be empty	New assumption (during the intervention)
dAA2[2]: If a UAS is (assumed to be) true, its converse *may* be true	New assumption (during the intervention)
aAAm2: If a UAS is (assumed to be) true, its converse *may* be false	New assumption (after the intervention)

I first describe Andrea's initial assumption about the relation between a UAS and its converse, and her related initial assumptions, then I comment on two concepts introduced during the intervention that Andrea used to make sense of one of her initial assumptions (bAA3). Finally, I focus on Andrea's reconsideration of her initial assumptions as well as the factors that seem to have influenced that shift and the way she manifested this change.

Andrea's initial related assumptions

Three of Andrea's initial assumptions about UASs: bAA1, bAA2 and bAA3, were observed during the First Exploratory Interview, before the intervention[24]. The teachers were requested to choose from a list of twelve statements those that conveyed the same thing as the *imaginary*[25] universal statement St2[26],

St2: *All «Vallejo» numbers are even numbers.*

Andrea picked (as expected) the four following statements:

- If a number is «Vallejo» then it is an even number (option c),
- Every «Vallejo» number is an even number (option g),
- The «Vallejo» numbers are even numbers (option h),
- There are no «Vallejo» numbers that are not even (option k).

However, Andrea was uncertain about choosing options "e" (*"The even numbers are not «Vallejo» numbers"*) and "j" (*"There are «Vallejo» numbers that are even numbers"*).

[23] The codes used for the teachers' assumptions have a specific meaning. For example, the code bAA3 means that it is a before-the-intervention ("b") assumption ("A") shown by Andrea ("A"), her third assumption ("3"). The code dAA2[2] means that it is a during-the-intervention ("d") assumption ("A") shown by Andrea ("A") during discussion 2 ("2"), her second assumption ("[2]"); whereas the code aAAm1 means that it is an after-the-intervention ("a") assumption ("A") shown by Andrea ("A") during a meeting ("m"), her first assumption ("1"). See Appendix "Notations" for other examples.

[24] For details on the task, as it was posed to the teachers, see Appendix EI1.

[25] In short, I call an *imaginary* statement to a statement whose truth value is impossible to determine. St2 is an example of an imaginary statement and its truth value cannot be determined because what "Vallejo" numbers are is unknown. For more details on "imaginary statements", see Chapter 4, Section II.2.1, Stage 1.3.

[26] Observe that I do not begin with St1 as I am following the numbering I used to list the statements that were included before, during and after the intervention. See Appendix Sts for the complete list of statements.

Andrea's explanation of why she hesitated about statement "e" exhibits her three initial assumptions:

> Andrea: *This (statement "e") is OK, partially… But all even numbers are not Vallejo [numbers]… Of course, because here it only says that all Vallejo [numbers] are even [numbers], but not all even numbers are Vallejo [numbers], … I understand that Vallejo [numbers] can be a group [she probably means a "set"] of numbers, but all this group of numbers are even. That is, it is a group of all even numbers, the Vallejo numbers is a subgroup [she probably means a "subset"]. Then… the even numbers are not Vallejo numbers, because when speaking about the even numbers I could refer to another [number], that is outside this subgroup. Therefore, in that case it does hold that-, the even numbers, those outside, are not Vallejo numbers. I'd better choose "e".*

When Andrea states: *"I understand that Vallejo [numbers] can be a group [she probably means a "set"] of numbers, but all this group of numbers are even. That is, it is a group of all even numbers, the Vallejo numbers is a subgroup [she probably means a "subset"]"* she expresses her assumption bAA1: *"All X are Y"* can be represented as set *X* included in set *Y*. Her use of this assumption also revealed that Andrea spontaneously relied on a set-based approach to interpret a universal statement; that is, she used a representation for the relation of the sets involved in the given universal statement to make sense of the task[27]. In addition, Andrea made a claim that suggested that she seemed to have taken for granted the existence of even number(s) that were not "Vallejo" numbers (*"when speaking about the even numbers I could refer to another [number] that is outside this subgroup. Therefore, in that case it does hold that-, the even numbers, those outside, are not Vallejo numbers"*). If we think of the UAS *"All X are Y"*, Andrea assumed that there were elements in *Y* that were not elements in *X*, or equivalently, there were elements in $Y - X$. This is supported also later during the intervention (in Discussion 2.4.2).

Andrea's assumption bAA2 (if a UAS is [assumed to be] true, its converse is false) is suggested by two aspects: (1) Andrea's focus on inferences, and (2) Andrea's denial of the converse statement. On one hand, Andrea's first claim showed that her attention was on an inference: that the given (assumed to be true) UAS *implied* that its converse was false (*"here it only says that all Vallejo [numbers] are even [numbers], but not all even numbers are Vallejo [numbers]"*). In concrete, her further reasoning of statement St2 suggests this, given that Andrea focused on the converse and whether it was false. In addition, her later use of "therefore" is also evidence of Andrea's focus on inferences she made from St2 (*"Therefore, in that case it does hold that-, the even numbers, those outside, are not Vallejo numbers. I'd better choose 'e'"*). Indeed, Andrea's puzzlement about whether or not option "e" was suitable stemmed from her focus on inferences that she considered viable. On the other hand, Andrea denied the possibility that the converse (*"All even numbers are «Vallejo» numbers"*) could be true based on St2 (*"but not all even numbers are Vallejo [numbers]"*).

Andrea's assumption bAA2 (if a UAS is [assumed to be] true, its converse is false) was general enough to not depend on a context, nor on agreement with common knowledge. This was shown by Andrea's use of the assumption for a statement that was in violation of practical knowledge. After Discussion 1.7 the teachers discussed the statement *"All dogs are dinosaurs"* (St39) and a representation for it.

[27] In fact, Andrea was the only teacher who used a set-based approach to make sense of a universal statement during the intervention without being requested to do so.

St39: *All dogs are dinosaurs*

Andrea's assertion, *"but not all the dinosaurs are dogs"*, exhibited the use of her assumption that if the UAS were assumed to be true, it would follow that the converse was false. Clearly, St39 was not a true statement, which did not prevent Andrea from inferring that if it were, its converse would be false.

Whenever the teachers encountered an imaginary statement like St2, they assumed that it was true, even though this condition was not mentioned or included in the task formulation. In fact, the teachers used that assumption when they were asked for the statement interpretation. Dawkins (2019) explained a similar phenomenon with one of Grice's everyday-conversation maxims[28] that might directly influence this assumption: "Do not say what you believe to be false". This maxim is used in daily-context conversations and according to Dawkins this might influence the way mathematical statements are understood. In this context, given that at least one of the involved classes (set of "Vallejo" numbers) in the imaginary statement St2 is not precisely defined, its truth value cannot be determined. However, as they were given the statement, they worked under an assumption of its truth value, which they assumed that was true; otherwise, they would have expected to be told that the statement was false.

Given that it is more common to expect that individuals assume that true UASs imply true converses, I did not anticipate Andrea's initial assumption bAA2. Mathematical reasoning is a specialized form of daily-life reasoning (Epp, 2003) and I see Andrea's initial assumption as her application of a combination of two of Grice's (1975, 1989) maxims of conversation to mathematics: *Maxim of quality* and *maxim of quantity*. While the maxim of quality involves not saying *false* things, the maxim of quantity refers to only saying what is really needed. Andrea might have assumed that since nothing was mentioned about the converse, it was false. This reminded me of an example given by Paenza (2010, p. 171) about a man who took an elevator and found two young girls therein. He turned to see one of the girls and told her: "You are very beautiful". The question Adrián Paenza posed was: Should the other girl feel less beautiful than the praised girl or not beautiful at all? Even though nothing was mentioned about the other girl, this is a very common interpretation of this kind of situation in informal, everyday contexts. But, logically, why should such a girl feel that way if she was not even mentioned in the man's comment? Think of the following additional example: A teacher said to her students "Every Wednesday during the semester there will be a math test". Now, here is my question: Does this mean that the students should expect a test *only* on Wednesdays? At least, this is what most of my students seem to have assumed through the years, based on their "in-shock" faces whenever I asked them to get ready for a test on a day different from Wednesday. They were strikingly surprised, though not happily. In these cases, the informal language common interpretation definitely played a role. It could explain why Andrea reasoned as she did. Based on her presumption that a UAS was true, she took for granted that the converse was not true because nothing was mentioned about it.

It is reasonable to assume that Andrea considered that a UAS and its converse did not state the same (her assumption bAA3) based on her consideration that those statements always had opposite truth values (her assumption bAA2). Clear evidence that Andrea used her assumption bAA3 was her decision not to choose options "b" (*"All even numbers are «Vallejo» numbers"*) and "i" (*"If a number is even then it is a «Vallejo» number"*), which are converses of the given statement. This means that even though Andrea's

[28] In Chapter 3, Section I.2.3, I included a complete version of Grice's maxims for everyday conversation.

assumption bAA3 was mathematically aligned, it is possible that the rationale she used to make sense of it was not. In that sense, Andrea's response is similar to the *type D* responses that were given by some of the high school students in Hoyles and Küchemann's (2002) study[29]. The truth value of the statement and the order of the elements in a statement (as I show below) seem to have played an important role in Andrea's consideration that a UAS and its converse did not state the same.

Andrea's incorporation of "Conditions" and "set of analysis[30]" to further explain her initial assumption bAA3

Two concepts that were introduced during the intervention played an important role in Andrea's further explanation of her initial assumption bAA3, that a UAS and its converse did not make the same claim; namely: "conditions" of a universal conditional statement and "set of analysis".

Different "conditions"

"Conditional statements" were introduced in Discussion 2[31]. The terms *antecedent* and *consequent* were also explicitly introduced then. *Antecedent* (the "if-condition") was introduced as the condition that was between "if" and "then" and *consequent* (the conclusion) as the condition that came after "then" in a "if-then-statement".

Discussion 2.4 was about universal conditional statements and their relation with their converse. The teachers were asked to individually read and reflect on Classroom Episode 13.2[32], prior to a whole-group discussion of it. Specifically, the teachers were expected to notice that in the classroom episode, a student, Carla makes the statement *"If a distribution is FWM[33], then the remainder is zero"* (St45), but the teacher begins her feedback by making reference to its converse (*"if the whole and fair distributions end up with a remainder zero, then they also hold the maximal condition"*). This reflected the teacher's own misunderstanding, which Andrea immediately identified.

> Andrea: *The teacher has distorted it. The teacher has changed it, because the condition is different. The condition given by Carla is that the remainder is zero, while the teacher says that the condition is that it is maximal.*

When Andrea said *"The condition given by Carla is that the remainder is zero"*, her attention was clearly on the "second condition" (the consequent or conclusion) in Carla's statement as she contrasted it with the "condition" given by the teacher (*"while the teacher says that the condition is that it is maximal"*). Thus, Andrea realized that the teacher in the classroom episode changed the condition of the original statement and that as such that would make a difference in the whole statement as a result. She used the term

[29] For details on the types of responses given by the high-school students in Hoyles and Küchemann's (2002) study, see Chapter 2, Section I.1.

[30] Recall that the "set of analysis" is made of the elements that satisfy the "antecedent" condition of the statement. For details see Chapter 3, Section II.1.

[31] See Appendix D2 for details on the content of Discussion 2. This includes the sub-discussions that are part of Discussion 2. For example, it includes the content of Discussion 2.4.

[32] Classroom Episode 13.2 was based on a real episode observed in one of the participant teachers' teaching during the 2017-intervention. For the content of the classroom episode, see Appendix CE13.

[33] The teachers used "FWM" as an abbreviation for the Fair, Whole and Maximal conditions of distributions, which refer to distributions where everyone gets the same number of objects, the objects are not broken, and the maximal number of objects as possible are distributed.

"condition" introduced in the intervention to explain her initial assumption that the UAS given and its converse did not state the same.

The "set of analysis" changes

Andrea also used the term "set of analysis" to explain why a UAS and its converse did not state the same. The term "set of analysis" was introduced right before Discussion 2, during a recap of Discussion 1; however, the concept began to be developed during Discussion 1.7 using other words[34].

In Discussion 2.4.1 the teachers were asked whether the conditional statement St45 and its converse stated the same.

St45: *If a distribution is FWM, then the remainder is zero*

Episode 1 shows the differing stances taken by Lizbeth and Andrea about the immediate equivalence between a conditional statement and its converse.

Episode 1

Turn	Who	What
1	Lizbeth	*Yes, only that it is the other way around. It is the same statement, but backwards.*
2	I	*Is it the same statement? Is it exactly the same statement, but the other way around? But, the question here is, do they state or communicate the same? Are they equivalent[35]? Do they convey the same thing? I mean, they are not exactly the same because there were some changes, but do they communicate the same information?*
3	Lizbeth	*Yes, I say yes.*
4	Andrea	*No*
5	Lizbeth	*Why not?*
6	Andrea	*Because the set of analysis changes.*

First, Lizbeth noticed that there was a change in the order of the elements in the statement, presumably based on Andrea's observation during Discussion 2.4 (see above); however, from her viewpoint, this was not an issue; she concludes that both statement St45 and its converse state the same thing (turn 1). As I thought that Lizbeth had not understood the question I posed, I prompted the teachers to think about the same task but expressed in alternative ways (turn 2). Nonetheless, that did not change Lizbeth's initial answer (turn 3), but it did lead Andrea to reject Lizbeth's answer (turn 4). Andrea used the *set of analysis* concept to argue that the statements are not the same due to Lizbeth's demand for an explanation (turn 5). Andrea's argument rested on the fact that both statements had different *sets of analysis* (turn 6)[36].

In short, the reasons Andrea provided to explain her initial assumption bAA3 (A UAS and its converse do not state the same) had a focus on the structure of the *"if-then-statement"*. First, her attention was placed on the different consequents, and later, in Episode 1 her explanation relied on the different *sets of analysis* the statement and its converse had in order to explain that they did not state the same. This shows Andrea's

[34] For details on the content of the discussions that were part of Discussion 1, see Appendix D1.

[35] During the intervention, we used the expression "equivalent statements" as "statements that state/convey the same".

[36] After Episode 1, I introduced the statement *"If it is a person, then it is a mortal"* (St49) in order to support Lizbeth's understanding who, even after Andrea's explanation, still remained skeptical.

incorporation of concepts from the intervention into her set of initial assumptions about UASs.

Andrea's reconsideration of her initial assumption bAA2: A Cognitive Conflict

In this section I focus on Andrea's emergent reconsideration of her initial assumption bAA2, (that the converse of a true UAS is false). This was a result of a cognitive conflict provoked when true biconditional statements were introduced.

Andrea used her initial assumption bAA2 in Discussion 2.4.2. The teachers were asked whether the imaginary statement St51 stated the same as its converse:

St51: *If a number is even, then it is a «fast» number.*

Episode 2 includes parts of the dialogue we had, where Gessenia seemed hesitant about whether or not the converse was implied to be true and the rationale that Andrea used to explain her disagreement.

Episode 2

Turn	Who	What
1	Gessenia	*If a number is even, then it is a fast number. It means that all even numbers are fast numbers… Yes [they are equivalent], because it says, if a number is even, then it is a fast number. It means that all even numbers are fast numbers, doesn't it? Then, if a number is fast, it does not necessarily mean that it is even, or it does? Because-, … Yes [that is what is claimed in the statement], it is claimed that all even numbers are fast…*
2	Lizbeth	*No, because the-, hmm, the antecedent has changed, the space [set] of analysis... According to what you (she means Andrea) explained before, no, because the antecedent changes, the set I analyze changes.*
3	I	*Gessenia, what is your doubt?*
4	Gessenia	*Does that mean that there are numbers that are not even and that are fast?*
	…	
5	Andrea	*Or maybe we could say that the classroom is (she probably means "represents") the [set of] fast numbers, and hmm, we-, she (Andrea pointed to Gessenia) is the even numbers, but not necessarily a fast number is going to be even. I mean, as sets.*

In this episode Andrea leveraged our classroom setting to provide a set-based explanation in response to Gessenia's doubt about the possible existence of «fast» numbers outside the set of even numbers (turn 4). Andrea's explanation is an analogy, which consisted of a classroom where there were more people than just Gessenia. This is further evidence of Andrea's initial assumption bAA1, that given a UAS of the form *"All X are Y"*, one could certainly find elements outside *X* that belonged to *Y*. In terms of the statement being discussed, Andrea is explaining that there are «fast» numbers that are not even numbers (turn 5), consistent with her initial assumption bAA1.

Additionally, Andrea's use of the expression "not necessarily" could be misleading (turn 5) since her analogy itself supported the undisputed existence of «fast» numbers that were not even. Strictly speaking, when the expression "not necessarily a fast number is going to be even" is interpreted from a mathematical perspective, it suggests that there may or may not be «fast» numbers that are even; that is, there is no reason to presume that the converse will undoubtedly be true or false. However, it was not a rare occurrence that during the intervention the teachers used expressions that in mathematics meant something different from what they were trying to say (e.g., see Andrea's initial use of the term "counterexample" in Section I.2.2.2).

The conflicting example: The case of true biconditional statements

Andrea's assumption about the relationship between a UAS and its converse began to change during Discussion 2.4.3, when true biconditional statements were introduced. The discussion included the statement St58, its converse[37], and a discussion about whether both statements stated the same thing.

St58: *If the unit-digit of a number is 0, then the number is divisible by 10*

The teachers managed to represent each statement by using Euler diagrams[38]. St58 was represented as the set made of the numbers with the unit-digit equal to 0 (*T* henceforth) included in the set of numbers divisible by 10 (*D* from now on), whereas the second statement St59 (the converse of St58) was represented as *D* included in *T* (see Figure 12 and Figure 13, respectively).

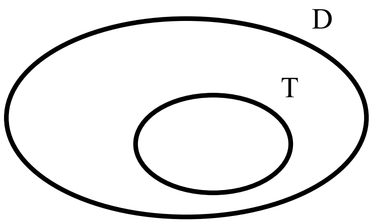


Figure 13. Diagram for the statement St58 "If the unit-digit of the number is 0, then the number is divisible by 10"

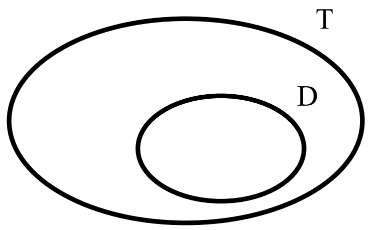


Figure 12. Diagram for the statement St59 "If the number is divisible by 10, then its unit-digit is 0"

The focus of the following episode, Episode 3 (below), is on Andrea's realization that there may be cases of true conditional statements for which its converse is also true and thus both main sets involved in the statement have exactly the same elements; namely, they are equal sets.

Once the two diagrams were represented, the teachers were asked if elements in the set *D* – *T* in the diagram in Figure 12 could be found (turn 1). Characteristics of elements in *D* – *T* were described (turns 4 and 6) and one possible number was eliminated as belonging to such a set (turns 7 and 8). This allowed Gessenia realize that *D* – *T* was empty (turns 8–10). As a result, Andrea noticed that *T* – *D* in the diagram in Figure 13 was empty as well (turns 11 and 12), which led her to conclude that both should be the same set (turn 14). This allowed the introduction of "biconditional statements", which was clearly a new concept for Andrea, based on her genuine surprise (turn 16). The teachers spontaneously suggested examples of what they believed were true biconditional statements. For instance, Andrea provided the example (St60), *"if s/he is Peruvian, then s/he was born in Peru"* (turn 16), which she presumably thought was a true biconditional statement. Based on my pointing out a possible exception to her example (turn 17), Andrea pointed out that *"It would not be a [true] biconditional, because there may be elements outside"* (turn

[37] The teachers became aware that both statements (the original statement and its converse) were true, even though we did not get into details about the respective proofs.

[38] A pictorial set-approach was first used during Discussion 1 of the intervention; however, more emphasis was placed during Discussion 2. The teachers were free to use a set-approach whenever it made sense to them in order to support their own understandings and as such they were not strictly required to use it every time. That meant that sometimes the teachers themselves (like Andrea) used it without being asked.

18). This shows that her attention was placed on the difference set, which she initially assumed was always non-empty (see her initial assumption bAA1, on which her assumption bAA2 is based).

Episode 3

Turn	Who	What
1	I	*And, can you find an element here (I point D – T in Figure 12 on the whiteboard)?*
2	Andrea	*No*
3	I	*What does this mean? What characteristics does an element in here (D – T in Figure 12) have?*
4	Andrea // Gessenia	*It does not finish in digit 0.*
5	I	*What else? How else should this element be?*
6	Andrea	*Divisible by 10.*
7	I	*… Could I have number 20 herein (D – T in Figure 12)?*
8	Gessenia	*No, because it should be inside [T]… There would be nothing over there [in D – T]*
9	I	*Then, this part over here (D – T in Figure 12), the one I'm coloring (I shade D – T on the whiteboard), how is it going to be?*
10	Gessenia	*Empty*
11	I	*… and over here (I point T – D in Figure 13 on the whiteboard), can I find elements?*
12	Andrea	*No*
13	I	*In other words, if this part (T – D in Figure 13) is empty, what does it mean in relation to the two sets?*
14	Andrea	*They are the same.*
15	I	*… both sets have exactly the same elements. That is the difference with the other examples we have seen. In the previous examples, we could find elements here, outside, right? But in this example, they are exactly the same sets… this is a biconditional statement.*
16	Andrea	*Ohhhh! … It is like saying, if s/he is Peruvian, then s/he was born in Peru.*
17	I	*The thing is, there are people who are Peruvians even though they were not born in Peru, for example people who got the Peruvian nationality. We should be careful with our examples in the daily-life context. Something similar happens with "if s/he was born in Peru, then she is Peruvian" …*
18	Andrea	*It would not be a [true] biconditional, because there may be elements outside.*

Unlike the example given (a *mathematical* statement), it is interesting that Andrea's example was framed in an everyday context. Andrea attempted to gather the characteristics she noticed in St58 (a true UCS with a true converse) in order to create her own example (St60: "*If s/he is Peruvian, then s/he was born in Peru"*); however, she did not anticipate that there were exceptions that made her example not completely suitable. My input (turn 17) was a warning about exceptions due to the ambiguity that the real-life context might create. Andrea realized why her statement St60 was not exactly an example of a true biconditional statement, because its representation *may* have elements outside the inner set, which differed from her initial assumption bAA1.

Andrea's responses to the conflict

Some studies show that individuals may not realize a conflict between their former assumptions and new conflicting data (e.g., Tirosh & Graeber, 1990); however, Andrea realized it and used her insight to modify her initial assumption. Andrea's responses to the conflict in Discussion 2.4.3 included: noticing the conflict with her initial assumptions, extracting key features of the conflicting example, creating her own conflicting example and modifying her initial assumption. Andrea's response to the conflict involves a *major change* to her set of initial assumptions and is similar to the "theory change" and "explicit knowledge building" types of responses reported by Chinn and Brewer (1993) and Chan, Burtis and Bereiter (1997), respectively[39].

Andrea's change of her initial assumptions was not only revealed through her refinement of her initial assumptions, but also through her construction of an example that illustrates her current insight. It exhibited the connections she made among her assumptions. Her example attempted to show a modification of her set of initial assumptions.

Factors that influenced Andrea's responses to the conflict

Chinn and Brewer (1993) pointed out some factors that may influence individuals' responses to anomalous data[40]. Among them, three seem to have had an effect in Andrea's responses: (a) *The individual's prior knowledge*. In particular, Andrea's mathematical background knowledge about divisibility supported her realization of the conflict. (b) *The anomalous data*. Notably, the contradicting nature of the conflicting example supported Andrea's change. Even though Andrea's initial assumptions were entrenched, this factor played a crucial role in Andrea's change. Particularly, the suitable sequence of questions (see Episode 3 above) drew Andrea's attention to the main contrasting aspects of the conflicting example in relation to her fundamental assumption bAA1. (c) *The processing strategies that guide the evaluation of the anomalous data*. Andrea's process of addressing the conflict was *deep*, in Chinn and Brewer's terms. Andrea was personally involved in the conflict. She was aware that the conflicting example was against her initial assumptions and wanted to resolve the conflict. She identified the main characteristics of the conflicting example that made her initial assumption invalid. Andrea's attempt to create her own conflicting example with those characteristics was her way to explain her new insights to others and establish her new assumptions. The expectation for explanations created in our context (a Proof-based teaching context) might have contributed to her deep processing of the conflict as it pushed Andrea to explain why she needed to modify her initial assumptions.

Further evidence of emergent change

Andrea's change was also observable after the intervention. During our Meeting #10 the teachers were engaged in a discussion about the imaginary statement St139 and five related tasks.

St139: *All natural numbers bigger than 5 are RAINBOW numbers*

[39] For details on types of responses to conflicting data, see Chapter 3, Section I.1.2.

[40] For details, see Chapter 2, Section III.1.

Specifically, in one task the teachers were asked if the given statement asserted that *the RAINBOW numbers were bigger than 5* (the converse of S139). Gessenia and Lizbeth said it did, but, Andrea considered that it was not the case.

> Andrea: *Because-, not necessarily, over there it says that all natural numbers bigger than 5 are Rainbow numbers, but there might be a natural number that is not bigger than 5 and that is Rainbow… 4 may be Rainbow.*

Even though her explanation exhibits Andrea's focus on inferences, this excerpt also shows that Andrea considered that given a UAS, it was uncertain whether the converse was false (it *may* be false). She considered the existence of the number 4 as an element outside the set of numbers bigger than 5 but possibly inside the set of rainbow numbers (*"there might be a natural number that is not bigger than 5 and that is Rainbow… 4 may be Rainbow"*). Moreover, it is interesting that she used her new assumptions with an imaginary statement, for which a specific truth value cannot be known.

Andrea's change was observable in terms of a subtle, but crucial, change in the language she used to express her new assumption. Her more accurate use of expressions, like *"there might be"* and *"may be"*, already suggested a shift from Andrea's initial assumption bAA2 about the falsehood of the converse statement (aAAm2: If A UAS is true, its converse *may* be false). I argue that this change was due to the cognitive conflict she experienced during Discussion 2.4.3, where Andrea became aware that there were cases of true conditional statements with true converses.

Andrea's emergent understanding

Andrea's emergent understanding that a presumably true UAS does not necessarily have a false converse was expressed in terms of her explicit modification of her fundamental assumption bAA1, which triggered the change of her related assumption bAA2. In particular, her change was expressed as her use of more *accurate language* to *refine* her initial assumption. Instead of suggesting that given a UAS, the converse will be false, Andrea was more cautious and her answer suggested a more precise expression (the converse *may* be false).

She used her identification of the main features of the conflicting example (a true UAS with a true converse) to try to create her own example (St60). This is evidence that her attention was now placed on those cases where her initial assumption did not hold anymore.

Andrea's case shows that the intervention had an impact on her initial assumptions about the relation between a UAS and its converse. Concepts such as "conditions in a statement" and "set of analysis", introduced during the intervention for teachers, reinforced Andrea's initial assumption that a UAS and its converse did not state the same (bAA3) by providing a rationale for why that was the case; while a discussion about true biconditional statements supported Andrea's realization that a true UAS does not necessarily imply that the converse is false, since it may be the case that the converse is also true.

Summary of Section I.1.1.1

Andrea's initial assumption about whether a UAS implies its converse differs from what has been broadly reported by other researchers (individuals tend to assume that based on a true UAS, its converse is also true, see Chapter 2, Section I.1). Instead, Andrea initially used the assumption that for a true UAS of the form *"All X are Y"*, its converse *"All Y are X"* is false (her assumption bAA2). This was based on a more fundamental initial assumption she used, that is: the UAS *"All X are Y"* can be represented as X included in Y with $Y - X \neq \emptyset$ (her assumption bAA1). Andrea also assumed that a UAS and its converse could not possibly state the same thing (her assumption bAA3) because of her other initial assumptions (bAA1 and bAA2).

During the intervention Andrea learned new concepts that supported her initial assumption bAA3; namely, the conditions (antecedent and consequent) in a UCS helped Andrea to explain that a UCS and its converse did not state the same, given that they had different conditions. In addition, the fact that the *set of analysis* in a UAS and its converse were in principle not the same supported Andrea's explanation of why both statements could not make the same claim.

Andrea's initial assumptions began to change as she learned about true biconditional statements during the intervention. A true mathematical biconditional statement provoked Andrea a cognitive conflict as it "broke" her reasoning pattern that true UASs always implied false converses. While her initial assumption bAA2 conclusively suggested the implication of a false converse, Andrea's emergent understanding left this implication only as a possibility and not as a necessary conclusion (her assumption aAAm2). Factors such as her mathematical background knowledge played an important role in this context. Her use of accurate language to refine her initial assumption bAA2 is evidence of her emergent understanding. Likewise, Andrea's attempt to create a new conflicting example suggested that she was aware of the conflict, but also that she did not reject it. Andrea resorted to the familiar context to come up with an example that complied the key characteristics of the conflicting example to illustrate her new insight.

Figure 14 shows the development of Andrea's assumptions about the relation between a UAS and its converse and the elements or aspects of the intervention that might have supported or triggered changes. The clouds include Andrea's assumptions. The first cloud includes Andrea's initial assumptions observed before, during or after the intervention. The last cloud includes Andrea's assumptions she ended the intervention with and were observed during or after the intervention. This includes new assumptions, but can also include initial assumptions that did not change with the intervention. The intermediate clouds have a focus on showing the assumptions that were used or changed. This means that it is possible to have a second cloud with only one assumption, even though the teacher had three initial assumptions. The arrows pointing left are inputs provided to the teachers, while the arrows pointing in the opposite direction are outputs that the teachers (here Andrea) provided and are relevant in their development. The outputs are in blue, whereas the inputs could be green or red. The green inputs did not trigger a change in the teacher's assumptions, while the red inputs might have triggered some sort of change/modification of the teacher's assumptions. The hexagonal boxes to the right of the inputs (or outputs) include the role the input (or output) played at that moment and for that teacher.

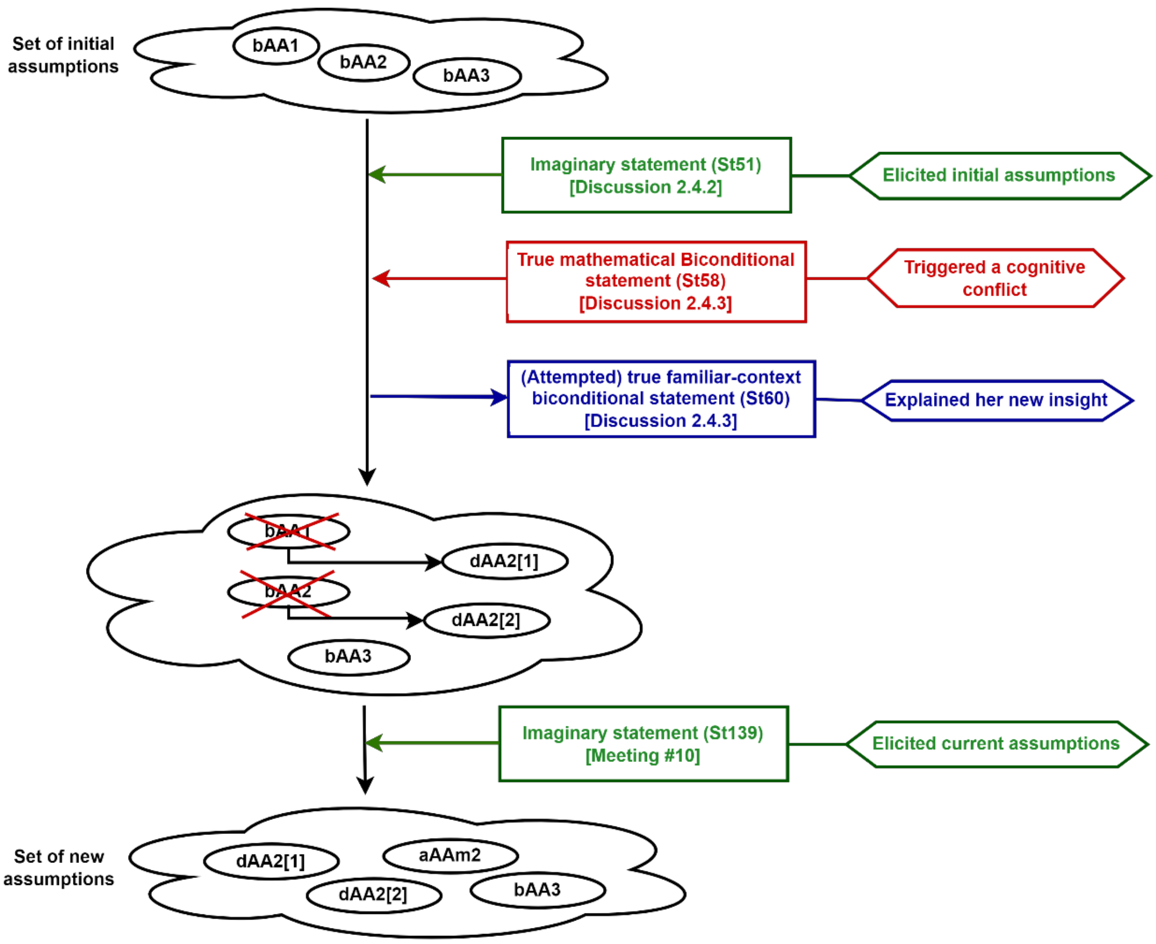


Figure 14. The development of Andrea's assumptions about the relation between a UAS and its converse (bAA1: "All X are Y" can be represented as set included in set Y, with Y-X≠∅; bAA2: If a UAS is (assumed to be) true, its converse is false; bAA3: A UAS and its converse do not state the same; dAA2[1]: "All X are Y" can be represented as X included in Y and Y − X may be empty; dAA2[2]: If a UAS is (assumed to be) true, its converse may be true; aAAm2: If a UAS is (assumed to be) true, its converse may be false)

1.1.2. Gessenia's assumptions

This section is about the development of Gessenia's assumptions for the relation between a UAS and its converse. This involves Gessenia's emergent reconsideration of her initial resistant-to-change assumption that a UAS and its converse stated the same, or as she initially expressed it as a UAS implied its converse, and vice versa.

Table 5 includes Gessenia's initial and new assumptions about the relation between a UAS and its converse. While her first two initial assumptions were identified before the intervention, her third initial assumption was only observable during the intervention. The three assumptions are interrelated and changed together.

Table 5. Gessenia's assumptions about the relation between a UAS and its converse

Assumption	Type of assumption (when first observed)
bAG1: If a UAS is (assumed to be) true, its converse is true	Initial assumption (before the intervention)
bAG2: A UAS and its converse state the same	Initial assumption (before the intervention)
dAG2[2]: In the UAS *"All X are Y"*, *X* and *Y* have the same elements ($X = Y$)	Initial assumption (during the intervention)
dAG2[1]: St49 (a specific UCS) and its converse do not state the same	New assumption (during the intervention)
dAG2[3]: In the UAS *"All X are Y"*, *X* and *Y* may not have the same elements	New assumption (during the intervention)
aAGm1: A true UAS does not necessarily imply that its converse is true	New assumption (after the intervention)
aAGm2: A UAS does not assert its converse, but it does not deny it either	New assumption (after the intervention)

I first present Gessenia's initial assumptions about the relation between a UAS and its converse; then I focus on Gessenia's reconsideration of her initial assumptions, which includes the factors that seem to have had an impact on them.

Gessenia's initial related assumptions

Gessenia's initial assumptions bAG1 (If a UAS is (assumed to be) true, its converse is true) and bAG2 (A UAS and its converse state the same) were identified during the First Exploratory Interview, before the intervention for teachers. The teachers were requested to choose from a list of twelve statements those that conveyed the same thing as the universal affirmative statement St2:

St2: *All «Vallejo» numbers are even numbers*

Gessenia chose the six following statements:

- *"If a number is «Vallejo» then it is an even number"* (option c),
- *"Every «Vallejo» number is an even number"* (option g),
- *"The «Vallejo» numbers are even numbers"* (option h),
- *"There are no «Vallejo» numbers that are not even"* (option k),
- *"All even numbers are «Vallejo» numbers"* (option b), and
- *"If a number is even then it is a «Vallejo» number"* (option i).

The first four statements selected by Gessenia were the ones the teachers were expected to choose. The fact that Gessenia (wrongly) picked options "b" and "i" suggested that

Gessenia assumed that the given statement and its converse conveyed the same thing (her initial assumption bAG2) since both "b" and "i" are the converse of St2, although presented differently ("b" as a universal statement and "i" as a conditional statement). She explained why she chose those options in the following terms, which confirmed my presumption.

> Gessenia: *Because if all Vallejo numbers are even numbers, then all even numbers are Vallejo numbers as well.*

Her explanation for why the statements stated the same (bAG2) was based on an equivalence she has taken for granted between St2 and its converse. Gessenia thought that once one of the statements is assumed to be true, its converse can be implied to be true (her initial assumption bAG1).

In the existing literature Gessenia's initial assumption bAG1 is known as *the converse error* and is very common when reasoning with universal conditional statements[41]. According to Epp (2003), students take for granted that the truth of one of the conditional statements implies the truth of its converse, unless their practical or daily-life knowledge goes against this assumption. In this context, it was impossible that practical knowledge had any influence in Gessenia's decisions, given that St2 is an imaginary statement. This shows that Gessenia's initial assumption bAG1 was independent of the context. It was general enough to be used in an imaginary context.

Furthermore, Gessenia's choices not only indicate the use of her assumption bAG1, but also that she rejected options that were existential statements (options "a", "f", "j" and "l"). This might suggest that Gessenia ignored the options that did not share the same universal *atmosphere* as St2 (Woodworth & Sells[42], 1935); that is, she overlooked the options that were not universal statements.

A potential conflict

The resistant-to-change nature of Gessenia's initial assumptions bAG1 and bAG2 was observed on several occasions during the intervention. For instance, in Discussion 2.3 the teachers were requested to write down a conditional statement that stated exactly the same as the statement St37,

> St37[43]: *All numbers divisible by 6 are divisible by 3.*

Lizbeth and Andrea gave correct equivalent conditional statements. After that Gessenia provided the statement *"If all numbers are divisible by 3, then they will be divisible by 6"*, which is the converse of the original statement St37. It showed again her use of her initial assumption bAG2 (A UAS and its converse state the same).

Gessenia's response suggests that she did not perceive any contradiction between her related assumptions bAG1 and bAG2 when applied to the case of St37. Observe that St37 is a true UAS with a false converse, so St37 and its converse could not state the same. In

---

[41] For details and references on *the converse error*, see Chapter 2, Section I.1.

[42] Woodworth & Sells (1935) explain that "[t]he atmosphere of the premises may be affirmative or negative, universal or particular. Whatever it is, according to the hypothesis, it creates a sense of validity for the corresponding conclusion" (p. 452). For details, see Chapter 2, Section I.1.

[43] The same statement was the focus of Discussion 1.7, where the main topic was universal affirmative statements.

Piaget's terms, Gessenia's was an *unadapted response*[44] given that she did not even realize the conflict.

This is an example of a *potential conflict* (Zazkis & Chernoff, 2008) that did not develop as a cognitive conflict. Three reasons may explain why that was the case: (a) It is possible that Gessenia's attention was not directed towards the truth values of the statements St37 and its converse. (b) Even if her attention was on the truth values, she might have been unsure of them. It is possible that her mathematical background knowledge was insufficient to be convinced of the truth values the first time they were discussed, in Discussion 1.7. (c) Her initial assumptions about universal statements (e.g., bAG2) pushed her to provide a quick response without considering other factors such as the truth value.

Given that St37 did not trigger a cognitive conflict, it failed as a *pivotal example*. Zazkis and Chernoff (2008) explain that "examples that 'fail' to serve as pivotal are examples that belong to conventional example spaces as understood by experts, but fall outside, at least temporarily, of personal potential example space of a learner" (p. 206). In Gessenia's case, I presumed that St37 belonged to both the conventional example space of mathematics as well as Gessenia's personal potential example space (Watson & Mason, 2005, p. 76). The first section of the intervention focused on the mathematical content (divisibility). Furthermore, Discussion 1.7 of the second part of the intervention included the same statement St37, its converse, as well as a debate about their truth value. These two experiences support my claim that St37 belonged to Gessenia's personal potential example space. However, it is possible either that she was not convinced of those truth values and as such she did not rely on them, or that her attention was not really placed on that issue and she mechanically used her initial assumption bAG2.

Gessenia's emergent understanding supported by context and practical knowledge

Identifying the set of analysis

After Discussion 2.4.1 (see Episode 1 above) the teachers were asked to determine whether a familiar conditional statement stated the same as its converse. The statement was:

St49: *If it is a person, then it is a mortal.*

The goal of adding this example to the discussion was to leverage Andrea's emphasis on the *sets of analysis* to explain that a UAS and its converse did not state the same (for details, see Section I.1.1.1 above). Moreover, the familiar context would facilitate the teachers' identification of the different sets of analysis for St49 and its converse.

Lizbeth gave a negative answer (St49 and its converse did not state the same) and Gessenia's response (*"it could be an animal"*) suggested that she supported her colleague's answer. Unlike previous cases, here Gessenia did not immediately assume that a conditional statement and its converse stated the same (her assumption bAG2). On the contrary, to support Lizbeth's negative answer Gessenia placed her attention on an element in the set of analysis of the converse statement (an animal is a mortal) that was not an element of the original statement St49. In fact, Gessenia's example proved that the converse of St49 was false (an animal is a mortal and is not a person).

[44] For more details on the types of responses to conflicting information, see Chapter 3, Section I.1.2.

It would be premature to conclude that, at this point, Gessenia understood that, because of the different sets of analysis, the statements did not state the same thing, as Andrea did. Instead, Gessenia seems to have relied on the particular (familiar) context of the example, her background knowledge about it and the truth value of the statements (Gessenia was aware that St49 was true and that its converse was false). This means that her implicit new assumption dAG2[1] that St49 and its converse did not state the same thing possibly remained attached to the specific example without being generalized.

Furthermore, Gessenia did not seem to have grasped the notion of "set of analysis" when it was introduced during Discussion 1.7. This was an obstacle for her to make sense of Andrea's explanation for why a UAS and its converse did not state the same. Her limited grasp was observable after the discussion of statement St49 and its converse. By then, the discussion had returned to the main statement of Discussion 2.4.1, St45, and its converse:

St45: *If a distribution is FWM, then the remainder is zero*

However, the focus was now on two new topics: first, the elements in the sets of analysis, and second, the truth value of the statements.

Gessenia asked whether the set of analysis was always located to the front of the statement. Her question was an indication of her recent realization of a pattern to identify the set of analysis in a statement[45]. In this sense, Andrea's explanation for why a UAS and its converse did not state the same (the sets of analysis are not the same) might not have been significant to Gessenia when Andrea came up with it, since Gessenia still struggled to grasp what the set of analysis was. In Piagetian terms, Gessenia did not possess yet the competence to respond to the stimulus I provided during Discussion 2.3. It seems that her lack of understanding of the concept "set of analysis" prevented her from properly responding to the stimulus. On the other hand, this seems to have been more easily recognized by Gessenia once it was situated in a familiar context (for St49), with which Gessenia might have felt more comfortable. This suggests that the context contributed to Gessenia's emergent understanding of the relation between a UAS and its converse.

Prior literature suggests that using examples of statements whose common interpretation is in harmony with standard logic might support the introduction of logical reasoning principles (e.g., Epp, 2003; Durand-Guerrier, 2008). Epp (2003) claimed that this was helpful "to motivate students to accept the reasonableness of the principles of mathematical reasoning" (p. 895). Notably, there exists evidence that when the context of Wason's selection task[46] is changed to a more familiar one, students do much better on the task (see, e.g., Cosmides, 1989; Valiña & Martín, 2016). Therefore, it is possible that Gessenia's emergent understanding might have been supported by the familiar context, which was aligned with Gessenia's practical knowledge.

Two Cognitive Conflicts

The first conflict: The elements of the sets involved in a UAS and assumption dAG2[2]

In Discussion 2.4.2 the debate was centered on the imaginary statement St51:

St51: *If a number is even, then it is a «fast» number.*

[45] In the examples we analyzed until that point, the sets of analysis were all to the front of the statement.

[46] The Wason's selection task is a famous logic task about inferences with conditional statements (see Chapter 2, Section I.1 for details on some results of research developed about it).

The question was posed, whether St51 states that *if a number is «fast», then it is even* (the converse of St51).

Episode 2 (see Section I.1.1.1 above) shows, on one hand, that Gessenia was aware that St51 was about *all* even numbers (*"It means that all even numbers are fast numbers"*, turn 1). On another hand, Gessenia hesitated about whether or not the converse could be implied from St51 (*"Then, if a number is fast, it does not necessarily mean that it is even, or it does?"*, turn 1).

The episode reveals an implicit initial assumption Gessenia used about UASs; namely, her initial assumption dAG2[2].

dAG2[2]: In the statement *"All X are Y"*, *X* and *Y* have the same elements

Gessenia's question in turn 4 (*"Does that mean that there are numbers that are not even and that are fast?"*) indicates that she had begun to hesitate about her initial assumption dAG2[2]. This means that Gessenia treated the two sets involved in the UAS as equal. The possibility that there might exist an element that belonged to the set of analysis of the converse statement (a fast number) and did not satisfy its conclusion (and is not even) puzzled Gessenia. It would reveal an inconsistency with her initial assumptions bAG1 and bAG2 given that the existence of a fast number that is not even would show that the converse is false and that the sets of analysis are not the same. Based on her initial assumptions, Gessenia interpreted conditional statements as biconditional statements. Accordingly, statement St51 would have been interpreted as *"A number is even if and only if it is a «fast» number"*. This would explain why Gessenia considered a conditional statement necessarily implied its converse (her assumption bAG1), which is an assumption that has been reported in the literature (e.g., Durand-Guerrier et al., 2012; Epp, 2003; Hoyles & Küchemann, 2002). Gessenia's assumption can be "justified" by features of everyday language. People usually interpret *"if-then-statements"* as if they were *"if-and-only-if-statements"*. In the case of Gessenia, this might be the result of her assigning an identifying role to the verb *to be* in statements of the form *"All X are Y"*, interpreting the verb *to be* as an equality (Veel, 1999)[47].

Figure 15 and Figure 16 illustrate Gessenia's initial and new assumption about a UAS and the elements of the sets involved in it, respectively. Both show my interpretation of the way Gessenia experienced an internal conflict.

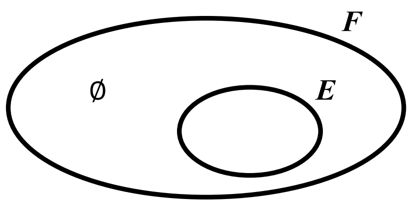


Figure 15. Gessenia's initial assumption about a UAS and the elements of the sets involved in it

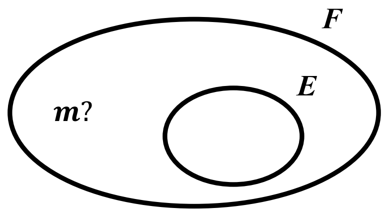


Figure 16. Gessenia's new assumption about a UAS and the elements of the sets involved in it

Episode 2 shows that Gessenia was conflicted by the differences between her initial assumption dAG2[2] and the possibility that it does not longer hold in general. It is likely that the discussion we had about St49 (*"If it is a person, then it is a mortal"*, see above) had introduced the conflict, but given that Gessenia's attention was placed on a different

[47] For details about some possible explanations why people might make the converse error, see Chapter 2, Section I.1.

aspect (the set of analysis), she might have not considered further consequences of it. When statement St51 was analyzed, doubt about whether her observations for St49 applied to the case of St51 surfaced, especially considering the imaginary nature of St51 (which made its truth value impossible to determine). Hence, St49 seems to have contributed to casting doubt on Gessenia's initial assumption dAG2[2]. It is possible that Gessenia had recalled what she had noticed in the discussion we had for St49 and she wanted to know whether what she noticed for such familiar-context example applied to the imaginary statement St51; that is, there may be elements that are not shared by both sets involved in a UAS.

Hesitation as a manifestation of Gessenia's emergent understanding

Gessenia did not conclude that given a "fast" number, it would certainly be even, as she might have done before because of her initial assumption bAG1. Thus, this is already an important step in the development of her understanding of the relation between a UAS and its converse. With the question Gessenia posed, she seems to have begun to reconsider her initial assumption for the elements in the sets involved in a UAS. In this particular case, her emergent change in her assumptions was observed as *hesitation about her initial assumption*.

The use of the imaginary statement St51 at this point (during Discussion 2.4.2) is pertinent. It served to find out whether Gessenia's initial assumptions had been altered at all as a result of the analysis of the previous familiar-context statement St49 during Discussion 2.4.1. Furthermore, given that the truth value of an imaginary statement is impossible to determine, I was confident that its truth value would not interfere when interpreting or making inferences about St51. The imaginary statement St51 supported my observations of the development of Gessenia's assumptions as it made her assumptions explicit. Additionally, it fostered Gessenia's own understanding as it pushed her to reflect on possible general consequences, which urged her to set aside the context as a determinant factor when reasoning her new assumptions.

I agreed with Gessenia's reaction, and I provided the new example St55 with similar characteristics to St49:

St55: *If a person is from Lima, then s/he is Peruvian*

This might have contributed to Gessenia's reconsideration of her assumption dAG2[2]. I presume that Gessenia became aware that there may be cases where the two sets involved in a UAS do not have exactly the same elements, which I call her new (implicit) assumption dAG2[3].

dAG2[3]: In the UAS *"All X are Y"*, *X* and *Y* may not have the same elements

Nonetheless, that does not seem to have been enough to change Gessenia's set of initial interrelated assumptions (bAG1, bAG2 and dAG2[2]). There is evidence for the persistence of her initial assumption bAG2 (a UAS and its converse state the same) as she applied it in the case of an imaginary statement during Meeting #10, after the intervention (this will be explained further in the next section, "Second conflict").

It is possible that Gessenia did not reconsider evaluating her complete set of initial assumptions because she had not established a clear linkage among her three initial assumptions, or her attention was only placed on one initial assumption (dAG2[2]) at the time. Analyzing St49/St55 had an immediate effect only on dAG2[2]. In that sense, her response is similar to Chinn and Brewer's (1993) *peripheral change* or Chan et al.'s

(1997) *surface-constructive* type of response as they involve minor changes to her set of initial related assumptions[48]. Gessenia still kept her assumption bAG2 (as I will show next), and possibly bAG1 as well. Notably, it seems that Gessenia's initial assumptions for the relation between a UAS and its converse were only weakly interrelated, given that once the modification of one assumption was contemplated, the modification of the other related assumptions did not immediately follow. Chinn and Brewer (1993) suggest that "it is important to identify those particular entrenched beliefs that are the source of the entrenchment of the theory as a whole" (p. 32). In this context, even though from an external perspective her set of initial assumptions is consistent, Gessenia does not seem to have been completely aware of the connections among them. Most likely, she needed more time to explicitly bring out those connections and further reflect on the effects of discussing St49 and St55 on them.

The second Conflict: Assumptions bAG1 and bAG2

A more concrete piece of evidence that the context played a relevant role in the development of Gessenia's understanding occurred after the intervention. During Meeting #10 the teachers were engaged in a discussion about the imaginary statement St139.

> St139: *All natural numbers bigger than 5 are RAINBOW numbers*

Specifically, one of the tasks asked the teachers to determine whether St139 states that *the RAINBOW numbers are bigger than 5* (i.e., the converse of St139). Gessenia's answer, unlike Andrea's, was affirmative. After Andrea's explanation (see Section I.1.1.1 above), I reminded the teachers of the familiar-context statement St55 that we had already discussed during the intervention for teachers (right after Discussion 2.4.2).

> St55: *If a person is from Lima, then s/he is Peruvian*[49]

This triggered the following response from Gessenia:

> Gessenia: *I get that, it [St139] does not assert this [its converse], but then it does not deny it [its converse] either… Of course, besides-, of course, it does not assert that [the converse], of course, because this is also as if I said, all the numbers-, um, all the numbers that go two by two bigger than 8 are even numbers, and I'm not necessarily saying that number 6, or number 4 are not [even]… I mean, if I understand that like this, of course.*

This is evidence that Gessenia's initial assumptions bAG1 (if a UAS is true, its converse is true) and bAG2 (a UAS and its converse state the same) have begun to change, as I explain next.

Evidence of change

Unlike in Discussion 2.4.2, here Gessenia did not express her thinking as a question with an expectation for confirmation. Instead, and presumably as a result of the familiar context of St55, Gessenia expressed the change of her initial assumptions in two ways:

---

[48] For details on the types of responses to conflicting information suggested by Chinn and Brewer's (1993) and Chan et al.'s (1997), see Chapter 3, Section I.1.2.

[49] The drawing of the map of Peru accompanied the statement, where Lima was included in it and pointed to it with an arrow.

(1) the refinement of her initial assumption and (2) her production of an example of a mathematical true UAS with a false converse.

First, Gessenia made a precise claim about the inconclusiveness of the truth value of the respective converse statement when assuming a UAS (*"it [St139] does not assert this [its converse], but then it does not deny it either"*). In general terms, Gessenia noticed that *a UAS does not assert its converse, but it does not deny it either* (her new assumption aAGm2), which is against her initial assumptions bAG1 (if a UAS is true, its converse is true) and bAG2 (a UAS and its converse state the same).

aAGm2: a UAS does not assert its converse, but it does not deny it either

Gessenia's new assumption clearly conflicts with her initial assumption bAG1. The rejection of bAG1 leads to a new assumption aAGm1:

aAGm1: A true UAS does not necessarily imply that its converse is true

The presence of aAGm1 is supported by the example she provided (*"all the numbers that go two by two bigger than 8 are even numbers"*), which is a true UAS with a false converse.

Second, Gessenia created a statement with the same key characteristics as those for St55 (a true UCS with a false converse), but in a mathematical context (*"all the numbers that go two by two bigger than 8 are even numbers"*). It is remarkable that Gessenia was able to produce an example with which she could illustrate her current assumptions about the relation between a UAS and its converse. The statement was framed in a mathematical context and she used it to explain that a true UAS does not necessarily imply that its converse is true. This means that at least at that stage Gessenia felt comfortable enough to produce a mathematical example with the same characteristics as the familiar-context statement used to trigger new insights about UASs and their converse.

It is uncertain whether the insight Gessenia gained about a UAS and its converse was about inferences and not about equivalences. Notice that the example she produced after the intervention is a true UAS and she showed that the converse did not necessarily follow; however, the original task referred to equivalences (whether or not statement St139 conveyed the same as its converse). Observe that Gessenia's initial answer was affirmative, but when the familiar-context example was brought up, she seemed confident that from a UAS, its converse cannot conclusively be inferred. Her answer (*"I get that"*) suggests this, but also the subsequent example she produced. In that sense, the familiar-context examples had an impact on Gessenia's understanding; however, her insight was more oriented to the inferences that may or may not be performed from a UAS and her awareness about equivalences may have been missed.

Factors that influenced Gessenia's responses to the conflicts

Statements St37, St49 and St55 share similar characteristics: they are true universal affirmative statements[50] with false converses. Notwithstanding, the context in which they are framed is different: St37 is a mathematical statement, while St49 and St55 are familiar-context statements. With St37 Gessenia did not perceive any contradiction as I mentioned before; however, with St49 and St55 Gessenia began to modify her initial

[50] In fact, St49 and St55 were UCSs, which involved that they could be written as UASs (see Chapter 3, Section II.2).

assumptions about the relation of a UAS with its converse. This suggests that the context played a crucial role in Gessenia's emerging understanding.

The familiar context of statements St49 and St55 and, more specifically, Gessenia's background knowledge related to them, played an important role in the development of Gessenia's understanding. Notably, St49 supported Gessenia's reflection on the "set of analysis" and how to identify it in a statement. Likewise, St55 supported the change of her initial assumption that a true UAS implied that its converse was true.

In addition, imaginary statements also played an important role. With the first imaginary statement she encountered (St2), it was possible to learn about Gessenia's initial assumptions. St51 and St139 supported Gessenia's further thinking about the observations she made based on the familiar-context statements St49 and St55, respectively. It allowed her to "jump" into general considerations about them that she might have immediately discarded before.

The imaginary context did not allow Gessenia to make a judgement based on the truth value of the statement. This favored her reflection and consideration of possibilities she had not previously thought of. For example, together with the concept "set of analysis" that she had begun to grasp (based on St49), analyzing St51 made Gessenia reconsider that the sets involved in a UAS had exactly the same elements, which she seems to have previously taken for granted.

Summary of Section I.1.1.2

Gessenia's initial assumption for the relation between a UAS and its converse was in tune with what has been usually reported by researchers as a typical misassumption some individuals make about converses[51]: a true UAS implies that its converse is true (her assumption bAG1). Gessenia's initial assumption bAG1 can be explained in terms of another assumption she used about the identification of sets *X* and *Y* in a UAS of the form *"All X are Y"*, which suggested her view of the non-existence of elements that were not *X* but were *Y*, if the UAS *"All X are Y"* was true (her assumption dAG2[2]). However, her assumption dAG2[2] was observed only later during the intervention.

Three conflicting examples were introduced during the intervention to support Gessenia's understanding. The first, a true mathematical UAS with a false converse (St37: *All numbers divisible by 6 are divisible by 3*), did not develop as a *pivotal* conflict given that it did not have a real impact in Gessenia's initial assumption that a UAS and its converse stated the same (bAG2). Presumably, Gessenia's weak mathematical background knowledge and her still evolving understanding of the concept "set of analysis" were obstacles for her to see the conflict with her initial assumptions.

The second example that triggered a cognitive conflict was an imaginary UCS (St51: *If a number is even, then it is a «fast» number*). Gessenia's conflict was expressed through her hesitation about her initial assumption dAG2[2], which was presumably stimulated by the familiar-context UCS St49 (*If it is a person, then it is a mortal*). St49, a familiar-context UCS with a false converse, seemed to have supported Gessenia's awareness of possible cases where her initial assumption did not hold. Gessenia's concrete traces of realization arose in the form of questions that she formulated as a way to anticipate possible scenarios she did not conceive as possible before: the existence of elements outside *X* but in *Y* based on the UAS *"All X are Y"*.

[51] For details on related literature, see Chapter 2, Section I.1.

The third conflicting example arose after the intervention. The teachers analyzed an imaginary UAS (St139: *All numbers bigger than 5 are RAINBOW numbers*) and whether it stated the same as its converse. Even though at first Gessenia agreed with it, the reminder of another familiar-context UCS (St55: *If a person is from Lima, then s/he is Peruvian*) was the stimulus that led her to conclude that the imaginary UAS did not assert its converse, and it did not deny it either (aAGm2). In order to further explain her insight, Gessenia produced an example of a true mathematical UCS with a false converse. With her example Gessenia exhibited her awareness that based on a true UAS, a true converse might not necessarily be inferred. Gessenia's assumption aAGm2 and her production of an example with characteristics that "broke" her initial assumption bAG2 are evidence of Gessenia's evolving understanding of the relation between a UAS and its converse.

Figure 17 shows the development of Gessenia's assumptions about the relation between a UAS and its converse and the elements or aspects of the intervention that might have supported or triggered changes[52].

[52] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

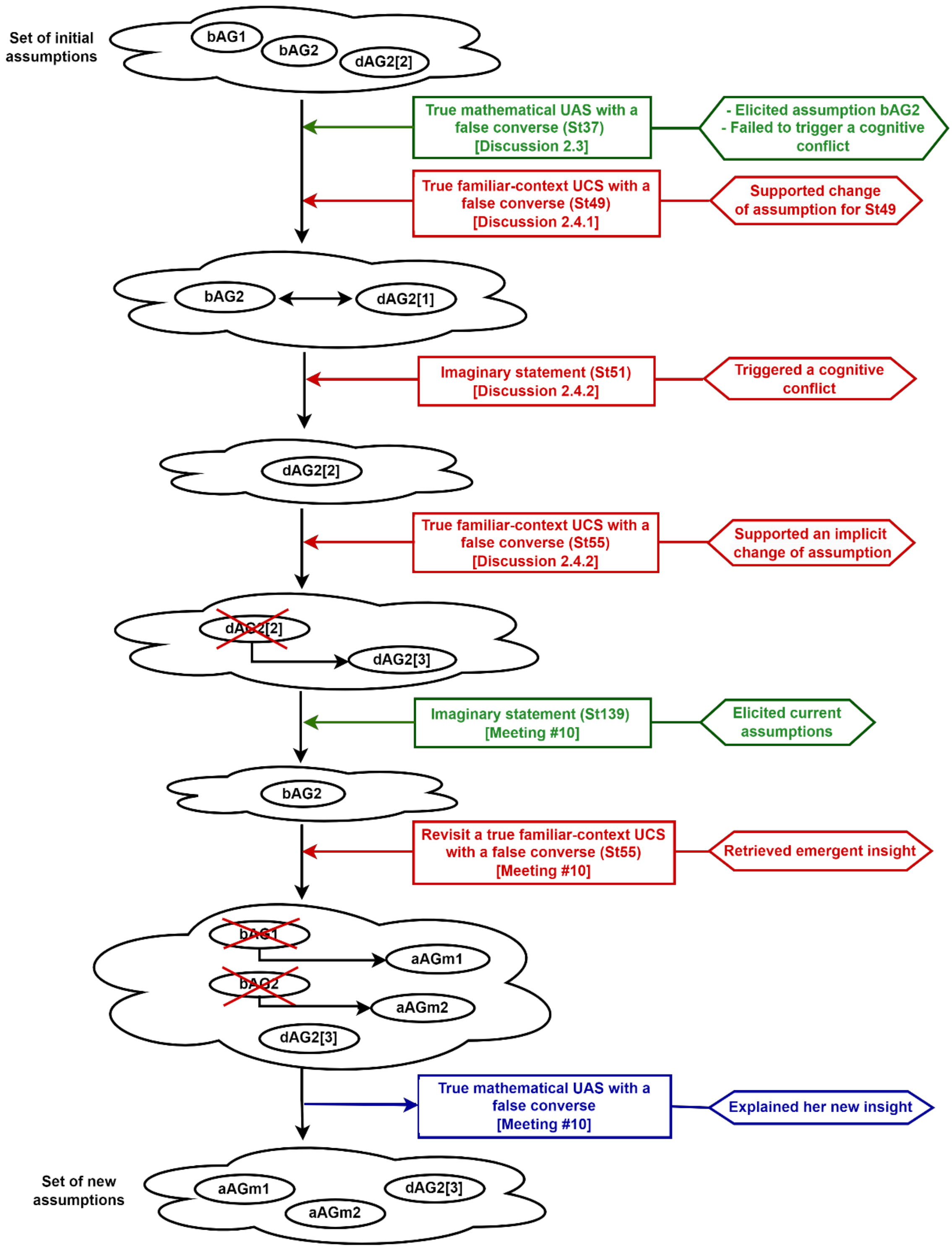


Figure 17. *The development of Gessenia's assumptions about the relation between a UAS and its converse (bAG1: If a UAS is (assumed to be) true, its converse is true; bAG2: A UAS and its converse state the same; dAG2[2]: In the UAS "All X are Y", X and Y have the same elements (X = Y); dAG2[1]: St49 (a specific UCS) and its converse do not state the same; dAG2[3]: In the UAS "All X are Y", X and Y may not have the same elements; aAGm1: A true UAS does not necessarily imply that its converse is true; aAGm2: A UAS does not assert its converse, but it does not deny it either)*

2. Disproving of Universal Affirmative Statements

In this section I focus on two main themes linked to disproving universal affirmative statements: the sufficient evidence to disprove a UAS (Section 2.1) and the description of all counterexamples that may disprove a UAS (Section 2.2).

A UAS might have more than one counterexample; however, from a mathematical perspective, one counterexample is sufficient evidence to disprove a UAS. A counterexample for a UAS of the form *"All X are Y"* would be an example that satisfies condition *X*, but does not satisfy condition *Y*.

I begin this section with an introduction to terminology I introduced to differentiate the types of counterexamples that I found that were important during the intervention and afterwards during my analysis. According to their existence[53] counterexamples can be classified in three categories: *possible*, *impossible* and *hypothetical* counterexamples.

A *possible counterexample* is a counterexample for a *false* universal statement. For instance, consider the false universal statement *"All even numbers are palindrome"*. A possible counterexample for it is the number 246 (an even number that is not a palindrome). It is called a "possible counterexample" because it is possible to construct one; that is, in theory it exists. In other words, the set of all the counterexamples for a false universal statement is a non-empty set and the elements in it are called *possible counterexamples.*

An *impossible counterexample* is a counterexample for a *true* universal statement. For example, in general, if *"All X are Y"* is a true universal statement, it follows that it does not have any counterexamples; however, it is still possible to talk about "impossible counterexamples" for it. They are non-existent examples that satisfy condition *X*, and do not satisfy condition *Y* of the statement. Given that the statement is true, the set of all counterexamples is empty. For instance, impossible counterexamples for the statement *"Every number divisible by 6 is divisible by 3"* are numbers that are divisible by 6 and are not divisible by 3, which do not exist.

A *hypothetical counterexample* is a counterexample for an imaginary statement (for details on what an imaginary statement is, see Section II.2.1, Stage 1.3, above). For instance, hypothetical counterexamples for the imaginary statement *"If a number is odd, then it is «poponupulus»"* are examples of odd numbers that are not «poponupulus». Theoretically, that is their characterization, even though their existence is only hypothetical.

In short, given a universal statement of the form *"All X are Y"*, a counterexample that would show that it is false have the same characteristics: it is an example of an *X* that is not *Y*. However, it is a *possible counterexample*, if it exists (the statement is false); it is an *impossible counterexample*, if it does not exist (the statement is true); it is a *hypothetical counterexample*, if *X* (the set of analysis) or *Y* are imaginary and its existence is hypothetical given that the truth value of the statement cannot be determined.

For the first theme (the sufficient evidence to disprove a UAS) I pay close attention to the development of Gessenia's and Andrea's understandings, their different initial expectations for what qualified as a mathematical justification to refute a UAS and the

[53] Recall that in Chapter 3, Section I.2.2, I adopted Peled and Zaslavsky's (1997) classification of counterexamples. Peled and Zaslavsky pointed out three types of counterexamples according to their level of sufficiency: specific, semi-general and general counterexamples. The classification used here is different, so any kinds of counterexample in Peled and Zaslavsky's classification could belong to any of my categories.

focus of the intervention on the logical interpretation of UAS that seems to have had a positive effect in their change of assumptions. For the second theme (the description of all counterexamples that may disprove a UAS) I include the development of all three teachers' understandings and the role that confirming and irrelevant examples played in their development.

I present Section 2.1 before the development of the identification of the characteristics for all possible counterexamples (Section 2.2) because one of the teachers (Andrea) produced a justification for such mode of argumentation at a very early stage of our proof-related discussions.

2.1. Sufficient evidence to disprove a UAS

In this section I focus on the development of Gessenia's and Andrea's assumptions related to sufficient evidence to disprove a UAS.

2.1.1. Gessenia's assumptions

Table 6 includes Gessenia's initial assumptions about the sufficient evidence required when disproving UASs, as well as the assumptions that she developed during the intervention and were observed during and after the intervention.

Table 6. Gessenia's initial and new assumptions about the sufficient evidence to disprove a UAS

Assumption	**Type of assumption (when first observed)**
dAG1[2]: The statement *"All X are Y"* is false because *not all X are Y*	Initial assumption (during the intervention)
dAG1[1][54]: Providing a "computation" is not sufficient justification to disprove a UAS.	New assumption (during the intervention)
aAGt1: I need to see; otherwise, it is not evident. Show me a counterexample.	New assumption (after the intervention)
aAGt3: When it says ALL and it is not true, it is sufficient to show a counterexample. There might be many counterexamples, but one is sufficient to disprove an "ALL-statement".	New assumption (after the intervention)

I present the development of Gessenia's assumptions by beginning with her (new) assumption about the insufficiency of a counterexample to refute a UAS, dAG1[1], instead of her initial acceptance of repetitive arguments when disproving UASs dAG1[2] because dAG1[1] was observed earlier than dAG1[2] during the intervention. Then I focus on her emerging understanding for which the logical interpretation of the UASs involved played an important role.

Gessenia's first related assumptions

Gessenia's expectation for a verbal justification (dAG1[1])

At the beginning of the second part of the intervention for teachers, Gessenia revealed her at-that-moment assumption in relation to the inclusion of a counterexample to disprove a UAS: Providing a "computation" is not sufficient justification to disprove a UAS (her assumption dAG1[1]).

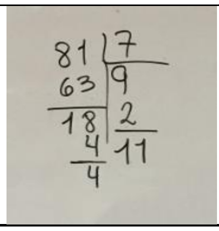
[54] Gessenia's assumption dAG1[1] was an assumption she developed during the first part of the intervention. This means that this is not exactly an initial assumption as she did not begin the intervention with this assumption.

Discussion 1.0 took place in order to identify, first, whether statement St16 was true and, second, whether the teachers (consciously or not) included a counterexample in their justifications to disprove St16:

St16: *All division of natural numbers are exact divisions*.

The teachers were asked to simultaneously solve the task individually on the whiteboard. Table 7 is a translation of the three teachers' written answers.

Table 7. The teachers' justifications that the UAS "All divisions of natural numbers are exact divisions" is false

Andrea's justification: It is false because when we divide a small number by a bigger number, the answer is 0, with the remainder equal to the small number. It means, it is a non-exact division.	**Gessenia's Justification:** It is false, because if the number of objects to be distributed is not enough to share out the same number of objects to each person, then we will have as the remainder a number different from zero. Therefore, it would not be an exact division.	**Lizbeth's Justification:** It is false because not all N [natural numbers] divided by N [natural numbers] are exact [divisions]. For example:  Result: 11 Remainder: 4

The three teachers agreed that the statement was false; however, only Lizbeth included a *specific*[55] *counterexample* (Peled & Zaslavsky, 1997) in her justification. On the other hand, Gessenia and Andrea provided *non-minimal justifications*[56]. Notably, they gave *semi-general counterexamples* (Peled & Zaslavsky, 1997).

It is worth mentioning that during the first part of the intervention counterexamples were not put forward as a topic and the teachers did not argue about false universal statements. This suggests that either Lizbeth unconsciously provided a counterexample (though she never called it as such), or she already counted on previous knowledge related to how to disprove universal affirmative statements as the one given here. The latter suggested a certain level of awareness that Lizbeth did not exhibit, though, as it will be seen in the discussion the teachers engaged afterwards.

Episode 4 (below) shows the discussion developed in relation to Lizbeth's justification (in Table 7) and specifically whether it was or not as a justification that disproved St16.

Gessenia's answer suggested that she believed that Lizbeth's justification did not qualify as a justification since it was not complete and she attributed its "incompleteness" to a lack of explanation (*"I think her justification is not complete. She does not say why."*, turn 4). What Gessenia missed in Lizbeth's justification was a narrative text to explain "why" (*"As we have always justified with text"*, turn 6).

Gessenia's main argument to reject Lizbeth's justification relied on its *form of expression* (A. J. Stylianides, 2007). In fact, this was a direct influence of the modes of representations of the arguments that were either produced or presented during the first part of the intervention. Given that the teachers mostly used arguments that were grounded on definitions, in our context that involved the use of text-form arguments. Gessenia linked the verbal form of a justification with the reasons to explain "why", and according to her, Lizbeth missed that in her justification. This shows that Gessenia had made an implicit assumption about the form of expression that mathematical justifications

[55] For some details on *specific, semi-general and general counterexamples* (Peled & Zaslavsky, 1997), see Chapter 3, Section I.2.2.
[56] For details on what I mean by non-minimal justifications, see Chapter 3, Section II.3.

should have. In addition, Gessenia revealed explicit uncertainty on whether a counterexample (the one used by Lizbeth), which she called an "exercise"[57], could be included in a justification (*"But can a justification include an exercise?"*, turn 6). She does not seem to have questioned the *mode of argumentation* (A. J. Stylianides, 2007) of Lizbeth's justification, but instead she argued whether a calculation could be included in a justification or not. Gessenia's focus on the form of expression led her to assume that a computation was, as a consequence, not a suitable justification (her assumption dAG1[1]).

Episode 4

Turn	Who	What
1	I	*Let's focus on Lizbeth's justification. Does it qualify as a justification? Is it sufficient?*
2	Andrea	*Hmm, here (she points to Lizbeth's justification) she is verifying it.*
3	I	*(I repeat what Andrea said) Here she is verifying it, hmm. Does that mean that this is not sufficient?*
4	Gessenia	*I think her justification is not complete. She does not say why.*
5	Lizbeth	*But I'm verifying it.*
6	Gessenia	*As we have always justified with text. But can a justification include an exercise?*
7	I	*It could be, but it depends. We will see on what exactly it depends later.*
8	Andrea	*I believe it [Lizbeth's justification] can be considered [a justification] because the statement says ALL, and if there is, at least one, as she wrote down, one that doesn't satisfy, then it is false. Since Pablito's statement[58] says ALL, then if there is some that doesn't satisfy, then it is no longer all.*

Previous research has also shown the influence of the mode of expression involved in an argument when its evaluation is in play (e.g., Barkai et al., 2002; Dreyfus, 2000; Knuth, 2002a)[59]. In Harel and Sowder's (1998) terms, Gessenia's initial scheme for disproving UASs was ritual given that the appearance of an argument, rather than its correctness, influenced the way she evaluated it. It shows that teachers' previous experiences with proving might significantly impact the assumptions they make, in particular those related to acceptable forms of expression for mathematical justifications. As I show in other cases (e.g., Section I.3.1.1 [Andrea's overgeneralization that no confirming example could prove a UAS] below), this was a general behavior the teachers showed during the intervention. They were prone to overgeneralizing patterns in the forms of reasoning they began to notice. Here Gessenia overgeneralized that justifications should have a textual form of expression.

Gessenia's acceptance of repetitive arguments (dAG1[2])

During the intervention Gessenia used another assumption to disprove UASs. It was an initial assumption that consisted in accepting that a sufficient justification to disprove a UAS of the form *"All X are Y"* could be the argument *"because not all X are Y"* (her assumption dAG1[2]). That is, Gessenia accepted an argument that did not include any mathematical evidence. Instead, it essentially repeated the statement with an added

[57] Gessenia used the word "exercise" to mean a numerical "calculation" or "computation". In this case it was the specific division of 81 by 7 that Lizbeth performed.

[58] Discussion 1.0 is contextualized as a statement given by a student named "Pablito" (for details, see Appendix D1).

[59] For details on these studies and other related ones, see Chapter 2, Section I.3.

negator in front of it. I call this kind of argument a *repetitive*[60] *argument*. In other words, Gessenia accepted an argument of the form: "the statement S is false, because S is not true".

In Discussion 1.1 the teachers were asked whether saying or writing *"because not all divisions of natural numbers are exact divisions"* was sufficient justification to disprove St16 (*"All divisions of natural numbers are exact divisions"*). This is a repetitive argument. Episode 5 shows the discussion developed about this issue.

Episode 5

Turn	Who	What
1	Andrea	*No, because it is repetitive.*
	Gessenia	*Yes*
	Lizbeth	*No*
2	Lizbeth	*It lacks an explanation. The justification is incomplete, maybe with a verification, that they verify why they say no.*
3	I	*Why do you say "yes" Gessenia?*
4	Gessenia	*Because it says it is false, because not all divisions of natural numbers are exact. Thus, it means, if not all are exact, then some-*
5	I	*Don't complete, just the way it was formulated.*
6	Gessenia	*Sure, yes, not all divisions are exact divisions.*
7	I	*The thing is, is it justifying something?*
8	Gessenia	*I mean, I, my justification would be-*
9	Lizbeth	*Do not complete!*
		(laughter)

While Andrea focused on explaining why the answer suggested in the discussion did not justify that the statement was false (*"No, because it is repetitive"*, turn 1), Lizbeth highlighted that there was something missing and pointed out what, in her view, it would be (*"It lacks an explanation. The justification is incomplete, maybe with a verification, that they verify why they say no"*, turn 2). Both teachers rejected the repetitive argument, and also provided reasons for why they did not agree with it.

In contrast, Gessenia explicitly agreed with the argument (turn 1). Gessenia's answer suggests that her agreement was based on her awareness that the statement was false[61] (see Discussion 1.0 above) and the fact that she could actually complete the argument they were given (*"Because it says it is false, because not all divisions of natural numbers are exact. Thus, it means, if not all are exact, then some-"*, turn 4; *"I mean, I, my justification would be-"*, turn 8). This might have led Gessenia not to see the need to include actual evidence to support a refutation.

Gessenia's answers in turns 4 and 8 included attempts she made to complete the argument by adding information that was not present in the argument they were given to analyze. This suggests that Gessenia was *"introducing personal knowledge"* (see Morris, 2007) to complete the given argument in order to evaluate it. Part of the knowledge she tried to introduce was directly related to her awareness that the statement was false. Morris pointed out that,

> to evaluate the soundness of an argument, a reasoner must "monitor the introduction of personal knowledge": the reasoner has to consciously set aside some prior knowledge and use only the premise information to determine what

[60] The term "repetitive" was inspired by Andrea' way to refer to this kind of arguments. See turn 1 in Episode 5.

[61] Recall that in Discussion 1.0 the teachers had disproved St16. See above.

> necessarily follows from the premises, while selectively introducing information from long-term memory to evaluate the truth of the premises. (p. 484)

In this case, Gessenia did not set aside her prior knowledge in order to evaluate the given (repetitive) argument, but rather she attempted to introduce the rationale of her current understanding.

Gessenia's answer suggested that dAG1[2] was an *initial assumption* that was not influenced in any form by Discussion 1.0, where the same statement was analyzed and specifically whether or not the inclusion of a counterexample yielded to a sufficient justification. Her initial assumption dAG1[2] shows that Gessenia was not aware that when evaluating the validity of an argument it is important to check the argument in its original form, without adding or removing information.

Later, during Discussion 3.2, Gessenia provided a repetitive argument of her own. Gessenia claimed that, *"because not all palindrome numbers are divisible by 11"*, statement St72 is false:

> St72: *All palindrome numbers are divisible by 11.*

Lizbeth immediately reacted by showing her expectation for a counterexample (*"and which one would be the counterexample?"*). Without any hesitation Gessenia provided number 131 as a counterexample, which suggested that she had already thought of such a case before Lizbeth requested it; however, her spontaneous response consisted in providing a repetitive argument. In this case Lizbeth played a crucial role in the development of Gessenia's understanding about the insufficiency of a repetitive argument when disproving a US. Lizbeth pushed Gessenia's awareness towards the need to provide a counterexample to refute St72.

My presumption is that after Lizbeth's input, Gessenia was more aware about the need to show a counterexample and this began to emerge as an implicit assumption. Additionally, as I show in Section 2.2 (see below), our discussions about the characterization for possible counterexamples also supported Gessenia's understanding of the sufficiency of one counterexample to refute USs.

Emerging understanding: Gessenia's teaching

Gessenia's demand for evidence: "I need to see"

After the intervention, during her teaching of Session 7, Gessenia displayed her focus on conclusive evidence by demanding the inclusion of a counterexample when disproving a UAS (her new assumption aAGt1). During that session and after the students solved the task, the teachers were expected to choose two or three groups' solutions in order to show them in front of the class and use them for a further whole-class discussion. The teachers needed to select answers with different, but relevant focus; for instance, a correct answer ("it is not true") with an incomplete argument, a correct answer with a complete argument, and an incorrect answer ("it is true"). The main goal of the teachers' teaching of Session 7 was to introduce the term *"counterexample"* since it was the first session where a false universal statement was discussed.

Gessenia asked her class to determine whether statement St45 was true or not and explain why.

> St45: *If a distribution is fair, whole and maximum (FWM), then there are zero objects left*

Gessenia's students worked in groups of three students and they were supposed to agree on an answer and write it down. For her whole-class feedback, Gessenia contrasted two different arguments, Argument A and Argument B. Argument A included a computation resulting in a counterexample, but Argument B was similar to the repetitive arguments Gessenia initially accepted as sufficient during the intervention (her initial assumption dAG1[2]).

> Gessenia: *They [the students who gave Argument B] could have used something that proved that what they claimed was true. I need to see, because what you tell me does not explain, it is not evident, I cannot see what you mean exactly. On the other hand, here (Gessenia pointed to Argument A), can I see what the group means? ... Yes, because they are showing me a fair, whole and maximal (FWM) distribution and it does not have zero left, therefore this [St45] is not true.*

Gessenia's explanation for why a repetitive argument was not suitable is now centered on the need for evidence (*"I need to see, because what you tell me does not explain, it is not evident, I cannot see what you mean exactly…"*). This already exhibits a deviation from her initial assumption dAG1[2]. The evidence Gessenia was missing in Argument B was a FWM-distribution that did not have zero objects left, that is, a counterexample. This was clear as Gessenia compared both groups' answers (*"On the other hand, here, can I see what the group means? ... Yes, because they are showing me a fair, whole and maximal [FWM] distribution and it does not have zero left, therefore this is not true"*).

This excerpt reveals not only a shift from Gessenia's initial assumption dAG1[2], but also a shift from her assumption dAG1[1], where she considered that a "computation" was not sufficient justification to refute a universal statement. Here she explicitly requested the group that provided Argument B to provide a counterexample, like that offered by the students who presented Argument A, implying her awareness that it was the evidence expected in that case.

Her change of assumptions dAG1[1] and dAG1[2] might be explained in terms of the extensive discussions we had during the intervention but also afterwards during our meetings. In all those episodes, Andrea and Lizbeth showed how they made progressive sense of two important aspects when disproving UASs: the sufficiency of one counterexample to disprove a UAS (see Section I.2.1.2) and the description for all counterexamples that disprove a UAS (see Section I.2.2). These two issues are important pillars for the disproving of universal statements and the fact that Gessenia gained these insights from the development of her colleagues' own insights seems to have been more meaningful[62] to her.

As a result, Gessenia's new assumption aAGt1 is evidence that her expectation for conclusive evidence to disprove a UAS evolved.

Gessenia's focus on the logical interpretation of UASs: The sufficiency of one counterexample

Another episode that exhibits Gessenia's emergent understanding of disproving UASs arose during her teaching of Session 10. Unlike her teaching of Session 7, Gessenia's insight is manifested through her emphasis on the interpretation of the UAS in discussion,

[62] I use the word "meaningful" to suggest that it was not imposed to her, but she could see the development of her colleagues' rationale for disproving UASs through our discussions and gain her own insights within this process.

which Gessenia used to explain the sufficiency of *one* counterexample to disprove a UAS (her new assumption aAGt3).

Gessenia's class was focused on statement St143 and on determining whether it was true and why.

> St143: *All natural numbers are divisible by 4,*

The following lines are part of the dialogue that Gessenia's class had.

> Gessenia: *It says ALL. Look what it says here, ALL natural numbers. Daniel, I have one question, how many natural numbers are there?*
>
> Daniel: *Infinite*
>
> Gessenia: *The natural numbers are infinite. Here it says ALL of them, ALL those infinite numbers, it says, are divisible by 4, and you all have said that this is not true. Here he [one student] gave [she refers to number 21], what? How have we called this?*
>
> Erin: *A counterexample*
>
> Gessenia: *A counterexample. He says, not, and how many examples did he need to show this?*
>
> Fatima: *One!*
>
> Gessenia: *One counterexample, the contrary. He is proving with a counterexample, just one. Miss, but there are many others. One-, one counterexample that breaks what is stated here is sufficient, ok?*

In order to explain statement St143, Gessenia emphasized the universal quantifier *"all"* (*"It says ALL. Look what it says here, ALL natural numbers"*) as well as the number of cases involved in the statement (*"The natural numbers are infinite. Here it says ALL of them, ALL those infinite numbers, it says, are divisible by 4"*). That is, Gessenia called her students' attention to the logical interpretation[63] of statement St143 before discussing its truth value. Gessenia brought out the possibility of finding more than one counterexample (*"Miss, but there are many others"*), but she emphasized the sufficiency of *one* counterexample despite the fact that other counterexamples could be found (*"One-, one counterexample that breaks what it is stated here is sufficient, ok?"*).

Figure 18 shows the development of Gessenia's assumptions about the sufficient evidence to disprove a UAS and the elements or aspects of the intervention that might have supported or triggered changes[64].

[63] For details on what I mean by the logical interpretation of a statement, see Chapter 3, Section II.1.

[64] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

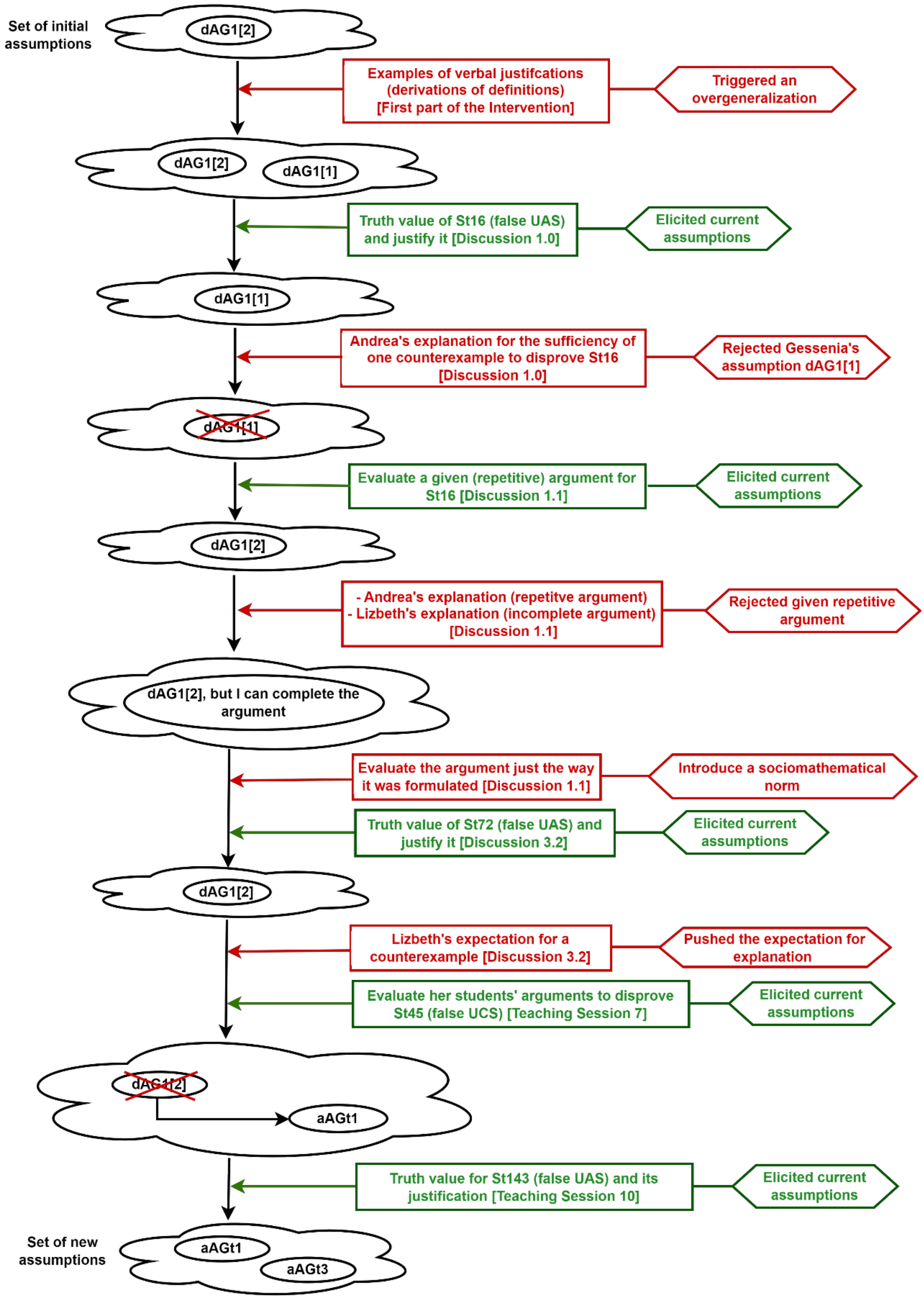


Figure 18. *The development of Gessenia's assumptions about the sufficient evidence to disprove a UAS (dAG1[2]: The statement "All X are Y" is false because not all X are Y; dAG1[1]: Providing a "computation" is not sufficient justification to disprove a UAS; aAGt1: I need to see; otherwise, it is not evident. Show me a counterexample; aAGt3: When it says ALL and it is not true, it is sufficient to show a counterexample. There might be many counterexamples, but one is sufficient to disprove an "ALL-statement")*

2.1.2. Andrea's assumptions

Like in the case of Gessenia, based on the first part of the intervention for teachers, Andrea first assumed that a counterexample (a "calculation", as she saw it) was not sufficient justification to disprove a UAS. Unlike Gessenia, Andrea's assumption changed quickly. Andrea's focus on the logical interpretation of the statement from the beginning played a crucial role on this shift from her first assumption.

Table 8 includes Andrea's assumptions about this issue, as well as the assumptions that she developed during the intervention and were observed during and after the intervention.

Table 8. Andrea's assumptions about the sufficient evidence to disprove a UAS

Assumption	**Type of assumption (when first observed)**
dAA1[2][65]: Providing a "computation" is not sufficient justification to disprove a UAS.	New assumption (during the intervention)
dAA1[3]: Providing a counterexample can be considered a justification because the statement says ALL and if there is at least one that does not satisfy, then it is false.	New assumption (during the intervention)
dAA1[6]: I show my counterexample and prove that not all.	New assumption (during the intervention)
aAAt2: When I say "all" or "none" and I give a counterexample to prove that it is false, that is valid.	New assumption (after the intervention)

In order to illustrate the development of Andrea's understanding about the sufficient evidence to disprove a UAS, I begin highlighting Andrea's initial rejection of example-based arguments, which entailed her assumption that a counterexample did not suffice to disprove a UAS; then I focus on the role that the interpretation of a UAS played to her emerging understanding, which was manifested in different ways that included refinements of some imprecise utterances she used during her teaching.

Andrea's first assumption

Andrea's rejection of example-based arguments (dAA1[2])

In Episode 4 (see previous Section I.2.1.1) Andrea exhibited her disagreement with Lizbeth's approach (see Table 7 above). According to Andrea it did not qualify as a justification that refuted the universal statement St16 in Discussion 1.0.

St16: *All divisions of natural numbers are exact divisions*

Two pieces of evidence explain Andrea's first approach to tackle the task I posed. First, Andrea did not include a counterexample in her own justification, which suggested that she might have found this either unnecessary or inappropriate. As I explained in the previous section this might have been influenced by the examples of justifications the teachers had seen until that moment. The statements previously analyzed were basically true universal statements and the justifications were mostly text-based derivations from definitions. Second, Andrea's answer (*"Hmm, here [she points to Lizbeth's justification] she is verifying it"*, turn 2 in Episode 4) was not a straight "yes" or "not" answer, which might suggest hesitation. Andrea's answer might be linked to a prior discussion traced back to the first part of the intervention (where the focus was on the mathematical content)

[65] Like Gessenia, Andrea developed her assumption dAA1[2] during the first part of the intervention. It implies that this is not an initial assumption. More details about it below.

in which Andrea asked about the difference between a *justification* and a *verification*. At that point I suggested that verification has to do with testing cases and that a justification, in the context of mathematics, involves more complex ideas. To the latter she reacted with the question, *"maybe something more theoretical?"*. I then asked her to wait for the upcoming discussions in order to exactly find that out on her own. I believe that as a result of that previous dialogue Andrea assumed that what Lizbeth did was *verifying* (in the sense of testing) and not *justifying* that the given statement was false.

It is interesting that, while Gessenia expected a specific form of expression when disproving a UAS (see previous section), Andrea expected a general "theoretical" argument instead. This is revealed by her own justification that is a mechanism to produce several (but not all possible) counterexamples (i.e., it was a *semi-general counterexample*, see Peled & Zaslavsky, 1997). Despite their different reasons, both teachers agreed on the insufficiency of Lizbeth's counterexample to disprove a UAS. None of these two teachers' justifications included a specific counterexample during Discussion 1.0 (see Table 7). Instead, both teachers' justifications involved a general property that we justified during the first part of the intervention.

A similar hesitation about the sufficiency of counterexamples to refute a universal statement has been shown as experienced by other elementary teachers in previous research (e.g., Barkai et al., 2002[66]). In Andrea's case, it stemmed from her previous experience with justification during the intervention[67]. Andrea was influenced by the first part of the intervention, which led her to assume that examples did not count as mathematical justifications and that was the main reason why she did not acknowledge Lizbeth's argument as a mathematical justification.

The role of the logical interpretation in determining sufficient evidence

Even though Andrea initially challenged the sufficiency of a counterexample to disprove a UAS during Discussion 1.0, her initial assumption shifted in the same discussion. I intended to open the possibility that "computations" could be included in mathematical justifications, but to let the teachers find out why on their own (turn 7 in Episode 4). This seems to have turned Andrea's attention to the universal quantifier "all" in the statement St16 and what it meant for St16 to be false.

> Andrea: *I believe it [Lizbeth's justification] can be considered [a justification] because the statement says ALL, and if there is, at least one, as she wrote down, one that does not satisfy, then it is false. Since Pablito's statement says ALL, then if there is some that does not satisfy, then it is no longer all.*
>
> (turn 8 in Episode 4)

Andrea explained why it made sense for her to have a "computation" in Lizbeth's justification. She pointed out two reasons for why that was the case: first, she emphasized that *"the statement says ALL"*, which shows her attention to the universal nature of the statement; second, Andrea highlighted that the existence of *"at least one… one that does not satisfy"* (even though she did not explicitly state yet what exactly it should not be

[66] For details and some other references, see Chapter 2, Section I.5.1.
[67] For more details on Andrea's previous experiences with justifications during the first part of the intervention, see Section I.3.1.1 (below).

satisfied[68]) made St16 a false statement. In other terms, Andrea grounded her answer on the logical interpretation of the statement.

In this sense, Episode 4 is interesting because it shows a quick change in Andrea's initial assumption about the sufficient justification that disproves a UAS, a shift from a *"no, a counterexample does not disprove the statement"* (her assumption dAA1[2]) to a *"yes, indeed it disproves the statement and these are my reasons…"* answer (her assumption dAA1[3]). By paying close attention to the meaning of the universal quantifier "all", Andrea found on her own the reasons for why St16 was false.

Certainty about the sufficiency of one counterexample (dAA1[6])

Andrea's firm certainty about the sufficiency of one counterexample to refute a universal statement was explicitly established later, during Discussion 1.4.2. Lizbeth was puzzled by St16 (*"All divisions of natural numbers are exact divisions"*) and its truth value given that it admitted many confirming cases. Lizbeth claimed that the confirming examples "proved" St16 and put forward the existence of many supporting examples as part of her argument[69]. Andrea's response to Lizbeth's conflict was conclusive: *"I show my counterexample and prove that not all"* (her assumption dAA1[6]). Andrea's answer revealed her certainty about the conclusiveness of a counterexample to reject an "all-statement". Furthermore, regardless of the many supporting examples that could be found, Andrea emphasized the sufficiency of one counterexample to show that the statement was false.

Andrea's Teaching

Imprecise language leading to inaccurate assumptions

Andrea's consistent use of her new assumptions dAA1[3] and dAA1[6] was clear throughout the intervention. An interesting situation emerged after the intervention, during Andrea's teaching of her Session 7. Two important aspects need to be the center of attention here: Andrea's focus on drawing attention to the logical interpretation of the statement in discussion and her use of inaccurate language, which resulted in some of her students' incorrect assumptions about the sufficient evidence to disprove a statement.

Episode 6 (below) shows extracts of the discussion triggered by the (false) statement St45 during Andrea's teaching.

> St45: *If a distribution is fair, whole and maximum (FWM), then there are zero objects left*

During her teaching Andrea drew her students' attention to the logical interpretation of the statement in discussion (turn 2). Note that Andrea emphasized that the given statement was a conditional statement and stressed its two main conditions, the antecedent and the consequent. More than that, Andrea highlighted that the first condition referred not only to one FWM-distribution, but *any* distribution that satisfied the condition of being FWM. Likewise, she brought attention to the implication of having those types of distributions according to the statement, namely, that there would *always* be zero objects left. Indeed, this was the approach she took as she walked around the students' desks minutes before the whole-class discussion, when they were solving the task in groups: Andrea asked

[68] In Section 2.2 I get back to this issue.

[69] For details on Lizbeth's conflict, see Episode 14 in Section I.3.1.2, below.

questions to the students in order to make sure that they understood the statement in the first place. Her goal, as it seems, was that her students understand why it made sense to include a counterexample as the sufficient evidence for the case in discussion. In that respect, it is interesting that she used the same strategy that worked well for herself as the basis of her own understanding during Discussion 1.0 of the intervention for teachers; namely, use the logical interpretation of the statement to identify what meant for it to be false.

Episode 6

Turn	Who	What
1	Andrea	*In order for their answer to be valid as a justification, they would need to include an example. Or, in this case, if they say NO, this is not true, when I am against a statement, it is called a COUNTEREXAMPLE. When I include a counterexample for the statement, then my justification is going to be valid.*
		(…)
2	Andrea	*Now, let's understand the statement. It says, if a distribution is fair, whole and maximal, THEN, when I am given this "then", it means that if this condition is satisfied, if it is any of these FWM distributions, if it is this type of distribution, it says, then there will be zero objects left; that means, then there will ALWAYS be zero objects left. Is that what is claimed in the statement? Does the statement claim that when the distribution is FWM, then there will always be zero objects left? Is that true? There will always be zero objects left?*
		(…)
3	Student A	*The remainder will not always be zero, sometimes there will be three left, or one, it depends on the number of people and the number of objects.*
4	Andrea	*Then give me an example of a FWM-distribution with a remainder different from zero.*
5	Student B	*Three marbles between two people, one is left.*
6	Andrea	*Does this example BREAK, BREAK this statement?*
7	Students	*Yes*
8	Andrea	*Then, when I want to prove that a statement is false, then I have to provide a counterexample, OK?*

Andrea's teaching aimed at explicitly pointing out the need to include an example (or as she refined it, a counterexample) in order to have a valid justification that disproved St45. This she did gradually in three steps: first, Andrea explicitly emphasized the need to include an example that would be against the statement, which was motivated by the lack of counterexamples in the groups' answers (turn 1); second, Andrea requested an example with specific characteristics (*"a FWM distribution with a remainder different from zero"*, turn 4); third, Andrea concluded the analysis of the statement reinforcing the need to include a counterexample when a false statement is under discussion (*"when I want to prove that a statement is false, then I have to provide a counterexample",* turn 8), even though she did not specify the kind of statement she referred to (whether it was universal or existential). Her last remark seems to have suggested to some of her students the sufficiency of one counterexample in order to refute *any* false statement as it is evidenced later during her teaching of Session 10.

It does not mean that during her teaching of Session 7 Andrea was not aware that in order to disprove a statement, the statement should not only be false, but also universal. During

Discussion 3.2 of the intervention, but also much later during one Meeting #9, she made it clear that she was aware of this condition[70].

> Andrea: *I believe examples are valid when the given statement is false and universal.*
>
> (Discussion 3.2)

Nevertheless, her teaching of Session 7 constitutes an interesting case as it shows that Andrea was not careful enough with her utterances as she omitted the type of the statement in discussion and, as a result, one of her students was misled to assume that it applied to all sorts of statements. The use of this assumption was evident during her teaching of Session 10, where her class analyzed the existential statement St142.

> St142: *Some natural numbers are divisible by 4*

What follows is an extract of the discussion.

> Student C: *If you say your answer is "not" [it is not true], then you should give a counterexample. You [Andrea] always give a counterexample when the answer is "not".*
>
> Andrea: *When it is false, but when we talk about "all", okay? or "none", ... when they are universal [statements]. When I say "all" or "none" and I give a counterexample to prove that it is false, that is valid.*

Andrea's refinement of the student's assertion exhibits her own insights about false universal statements and what entails to disprove them. Unlike the claims she used during her teaching of Session 7, in her teaching of Session 10 she explicitly emphasized that the statement should be not only false but also an "all" or "none" statement in order to use a counterexample to disprove it (her assumption aAAt2). With her response Andrea showed her consistency with her assumption that counterexamples are sufficient evidence to disprove universal statements. The need for explicitness about the cases where a counterexample was suitable was evident when she taught her own sessions, despite that her students were not familiarized with the terminology she used in her feedback (e.g., "universal statements" and "valid"). This I pointed to her, as I accompanied Andrea during her teaching, and she took such observation into consideration in her teaching as I show next.

Drawing attention to the logical interpretation of the statement in order to overcome a misassumption: The set of analysis

The next task in Andrea's teaching Session 10 involved the analysis of the universal statement St143.

> St143: *All natural numbers are divisible by 4*

The students were asked to identify whether St143 was true and justify their answers. Here Andrea explained what she meant by universal statements, she emphasized the interpretation of St143 and explicitly highlighted the sufficiency of one counterexample to disprove it.

[70] Nonetheless, at some point of Discussion 7 about the falsity and disproving of existential statements Andrea assumed that a counterexample would disprove an ES, she later changed her assumption (for details, see Section II.2).

Andrea: *Kids, something he [student C] was telling us before and that you should always remember, when I have a statement like this one, and it is universal, universal because it begins with "all", because your set of analysis, what you are supposed to analyze are AAAAALL the natural numbers, when you have a set of analysis that refers to all, then that is a universal statement. When your universal statement is false-, wait, is this true or false?*

Students: *False*

Andrea: *What is it sufficient so that you justify?*

Student C: *One counterexample*

Andrea: *One counterexample, and here [on the whiteboard] we have a counterexample. Is this sufficient in order to assert that this [St143] is false?*

Students: *Yes*

In order to explain why the statement in discussion was universal, Andrea used the concept "set of analysis", which was introduced during the intervention for teachers. The way Andrea instructed her class exhibits her own understanding of universal statements and how the set of analysis played a role in it. Unlike her teaching of Session 7, here Andrea was more explicit about the universality of the statement and its falsehood in order to use counterexamples as sufficient evidence to disprove them. Unlike the previous episode of her Session 10, the statement in this episode was a good opportunity for Andrea to get into details about what she understood by universal statements. Andrea's considerations were consistent with her assumptions dAA1[3], dAA1[6] and aAAt2 about the sufficiency of one counterexample to disprove a universal statement.

In general, the three teachers made similar explicit requests for evidence to their students during their teaching. Andrea expressed this as *"a counterexample needs to be included, then my justification is valid"* and *"you have to tell me which ones do not hold";* while Lizbeth did it as *"a counterexample is missing"*.

Figure 19 shows the development of Andrea's assumptions about the sufficient evidence to disprove a UAS and the elements or aspects of the intervention that might have supported or triggered changes[71].

[71] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

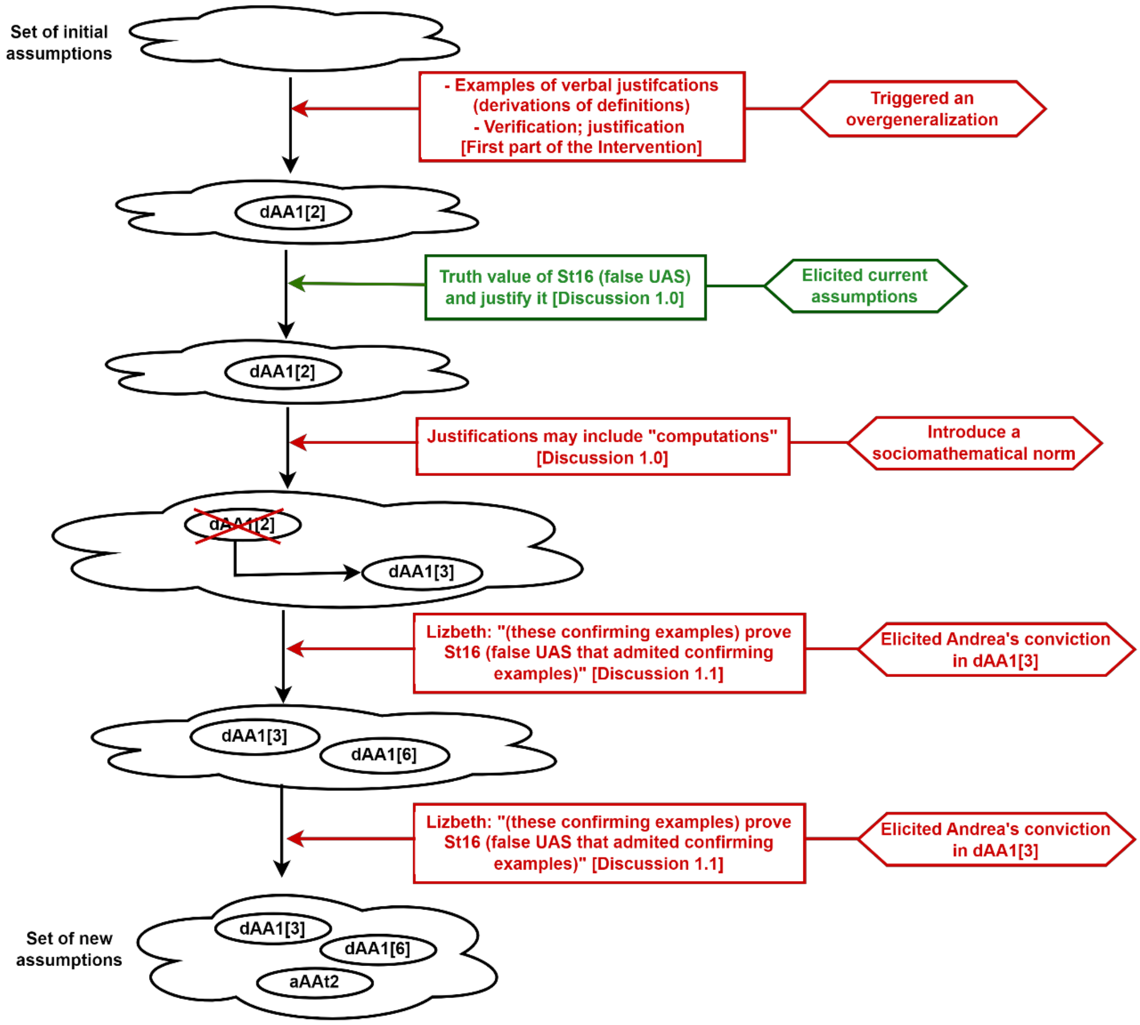


Figure 19. *The development of Andrea's assumptions about the sufficient evidence to disprove a UAS (dAA1[2]: Providing a "computation" is not sufficient justification to disprove a UAS; dAA1[3]: Providing a counterexample can be considered a justification because the statement says ALL and if there is at least one that does not satisfy, then it is false; dAA1[6]: I show my counterexample and prove that not all; aAAt2: When I say "all" or "none" and I give a counterexample to prove that it is false, that is valid)*

Summary of Section I.2.1

Assumptions like the rejection of the inclusion of a "calculation" (in the sense of a specific computation, but in fact a counterexample) in a justification were directly influenced by the first part of the intervention. At the beginning of the second part of the intervention Gessenia and Andrea used their assumption that a specific computation should not be included in a justification. The examples of text-based justifications the teachers had seen in the first of the intervention had an impact on the way the two teachers initially analyzed false UASs and the evidence they presented to disprove them; however, the teachers' attention was placed on different aspects. Whereas Gessenia was focused on the mode of expression of the argument (she expected a narrative argument), Andrea's focus was on the mode of argumentation (she expected a general-property argument).

The logical interpretation of statements played a crucial role to explain why a counterexample was sufficient to disprove a UAS during the intervention. Andrea's rationale was directly linked to the universal quantifier "all" in the given "all-statement", its meaning and the set of analysis.

Both teachers exhibited their new insights during their teaching, where they used similar approaches to the one that seems to have influenced their change of initial assumption during the intervention. Andrea and Gessenia resorted to the logical interpretation of statements to support their students' understanding of why one counterexample is sufficient evidence to disprove a UAS.

Additionally, in contrast to Andrea, Gessenia initially accepted a repetitive argument to disprove a UAS. Her lack of monitoring the introduction of her personal knowledge explained her behavior. Her colleagues' remarks seem to have had an impact on Gessenia's understanding, which was reflected as her criticism of a repetitive argument during her teaching.

2.2. Characteristics of counterexamples to a UAS

In general, a counterexample that refutes a UAS of the form *"All X are Y"* must *satisfy* condition X*, but contradict or deny* condition Y. In this section I focus on the factors that influenced the process of becoming aware of the characteristics of counterexamples that may disprove a UAS. This section complements Section I.2.1 in terms of understanding the disproving of UASs.

The only teacher that began the intervention with an initial usage of "counterexample" was Andrea; however, her use of the term did not match its mathematical meaning. During Discussion 1.2 I provided as an input an initial incomplete definition of "counterexample", that the teachers were explicitly requested to progressively refine: *a counterexample is a case that "breaks" the statement; it is against the statement*. I consider this initial input an implicit initial assumption that the teachers began the intervention with. The approach I used to develop the teachers' refinement of the characteristics of counterexamples was based on the discussion of *irrelevant* and *confirming* examples[72]. While confirming examples were expected to be rejected as a way to disprove a US because of the initial input, irrelevant examples were expected to be discarded as counterexamples because the logical interpretation of the statement requires focusing on the set of analysis X and its elements.

I focus on the development of each teacher's assumptions, beginning with Lizbeth.

2.2.1. Lizbeth's assumptions

Lizbeth did not use any initial assumptions about counterexamples, which suggested that this was a new mathematical concept for her. Table 9 shows Lizbeth's assumptions related to the characteristics that counterexamples have, which developed during and after the intervention.

[72] While an *irrelevant* example for "All X are Y" does not satisfy condition *X*, a *confirming* example satisfies both conditions *X* and *Y*. For details, see Chapter 3, Section I.2.2.

Table 9. Lizbeth's assumptions about the characteristics of counterexamples to a UAS

Assumption	**Type of assumption (when first observed)**
dAL1[1][73]: A counterexample "breaks" the statement	New assumption (during the intervention)
dAL1[2]: It (a confirming example) cannot be a counterexample to a UAS (St16) because it supports the statement	New assumption (during the intervention)
dAL1[4]: It (an irrelevant example) cannot be a counterexample to a UAS (St35) because it does not satisfy the first condition of the statement	New assumption (during the intervention)
dAL1[5]: The order in the statement is important	New assumption (after the intervention)

Lizbeth was the teacher who most actively participated during the discussions that aimed at determining a characterization for all im/possible counterexamples to UASs. In the following, I show that Lizbeth identified a way to discard challenging examples as counterexamples. Nonetheless, her reasoning did not seem completely general, but tied to the specific statements that we analyzed.

Lizbeth's rejection of confirming examples as counterexamples (dAL1[2])

Discussion 1.4.1 was a crucial point of the intervention in terms of whether certain examples qualified or not as counterexamples for a given statement and identifying characteristics of counterexamples. It had the focus on noticing that a confirming[74] example could not justify that a UAS was false. A classroom episode was used, in which students suggested different examples to falsify statement St16:

St16: *All divisions of natural numbers are exact divisions*.

In the classroom episode, Leo was the only student who suggested a confirming example: *5 divided by 1*[75]. The teachers were asked to determine whether Leo's example disproved St16[76]. Episode 7 (below) includes the teachers' answers and the main focus is on Lizbeth's explanation for why Leo's confirming example did not refute St16.

The teachers' immediate first response was affirmative (turn 1), perhaps directly influenced by their personal knowledge that St16 was false as it had been analyzed in previous discussions. This means that they might have initially been only focused on the truth value of the statement, without paying close attention to the characteristics of the suggested example. Previous research (e.g., Zaslavsky & Ron, 1998) has already shown evidence of students who accepted or used a confirming example as if it contradicted a given statement.

Andrea (incorrectly) pointed out that Leo's example was a counterexample (turn 6), and Lizbeth explicitly showed her disagreement (*"what Leo is providing is not a counterexample"*, turn 10) and explained that Leo's example was actually *confirming* the statement as it was the case of an exact division (*"It is affirming that it is exact"*, turn

[73] As I pointed out before, I provided an initial input for what a counterexample was and regarded dAL1[1] as that implicit assumption Lizbeth began the intervention with.

[74] In short, a confirming example is an example that satisfies the statement. During the intervention neither did I use the term "confirming" as a label, nor was it used by the teachers. Instead, we referred to them by their characteristics.

[75] For details on the content of the Discussion 1.4.1, see Appendix D1.

[76] Recall that in a previous Discussion 1.0 the teachers had already disproved St16 (see Section I.2.1 above for details).

10). That is, Lizbeth discarded Leo's example as disproving St16, as she noticed that it supported the statement instead of rejecting it as a counterexample would do.

Episode 7

Turn	Who	What
1	All the teachers	*Yes [5 divided by 1 disproves St16].*
2	Lizbeth	*Leo is giving his verification [---].*
3	I	*(I read the question again) Does Leo's example justify that the statement is false?*
4	Andrea	*No.*
5	Lizbeth	*But, in order to say that this is false; ah, then no.*
6	Andrea	*The thing is that it is not an example, it is a counterexample.*
7	Gessenia	*No.*
	Lizbeth	*I say no.*
8	Andrea	*Wait, yes.*
9	I	*No?*
10	Lizbeth	*Because, what Leo is providing is not a counterexample. It is affirming that it is exact.*
11	I	*It is not a counterexample, that's right.*
12	Andrea	*Ah, right!*
13	Gessenia	*On the contrary, it [5 divided by 1] supports the statement.*

In Lizbeth's answer (turn 10) it was implicit that she expected a counterexample to justify that the statement was false and it was explicit that she recognized that Leo's example was not a counterexample for St16, as she explained why. Lizbeth argued that Leo's example could not be a counterexample, presumably based on my initial input of what a counterexample was. My initial input suggested that a counterexample was a case that somehow "broke" a statement, that it was a case against a statement (assumption dAL1[1]); however, the suggested confirming example did not "break" St16 in any way. Instead, it supported the statement, as Gessenia pointed out (turn 13). Lizbeth's answer revealed not only her realization that an example that confirmed St16 could not disprove it, but also made clear her expectation that a counterexample is needed in order to refute St16. As a result of her engagement with this task, Lizbeth realized that confirming examples to a UAS could not be counterexamples to it, at least in the case of statement St16 (her assumption dAL1[2]).

Manifestation of her awareness: Lizbeth's Teaching

Lizbeth's use of her assumption that a confirming example to a UAS could not disprove the UAS (her assumption dAL1[2]) was consistent not only throughout the intervention, but also afterwards. For example, during her teaching of Session 7, Lizbeth's class discussed two groups' answers in relation to the truth value of statement St45 and its respective justification.

> St45: *If a distribution is fair, whole and maximal (FWM), then there are zero objects left*

Both groups agreed that the statement was false. Lizbeth introduced the term "counterexample" when she reviewed the first group's answer since the group included a counterexample in their justification. When providing feedback for the second group's answer, which did not include a counterexample, she asked her class whether adding the (confirming) example *3 divided by 3* would complete the group's justification. Lizbeth intended that her students could identify that an example was missing in such a justification, though not any kind of example, but a counterexample. In this sense, with the confirming example that Lizbeth introduced, she expected her students to reject it as the kind of example required to disprove St45. That is, she tried to guide her students to notice that a confirming example could not disprove St45; she wanted them to notice that her example was not a counterexample. For her teaching, Lizbeth used a similar approach as the one I used during the intervention for teachers, which revealed her own awareness that confirming examples cannot be counterexamples and disprove a statement.

Lizbeth's rejection of irrelevant examples as counterexamples (dAL1[4])

The main focus of Discussion 1.6.1. was a description for all possible counterexamples to St35[77].

St35: *The people present in this classroom are minors*

In particular, the teachers were explicitly requested to identify the characteristics of the counterexamples that justified that St35 was false. Episode 8 (below) shows the dialogue developed while we discussed the task. Its main focus is on two challenging examples ("my father" and "a 5-year-old boy in this classroom")[78] that I suggested as the discussion took place.

Gessenia suggested that a counterexample should not satisfy at least one condition in the statement (turn 1, see Section I.2.3 below for more details about Gessenia's assumptions). Lizbeth's answer, on the other hand, implied that the second condition of the statement should not be satisfied (turn 3). Motivated by Lizbeth's answer and aiming at getting a more precise description (explicitly that a counterexample should also satisfy the first condition of the statement), an example that did not apply to the situation in discussion, an *irrelevant example*[79], was put forward. The teachers were asked to determine whether the example *"my father"* qualified as a counterexample for St35 (turn 4). The suggested example was aligned with both Gessenia's and Lizbeth's ideas: it did not satisfy at least one condition in the statement and it did not satisfy the second condition "being a minor". Lizbeth rejected the irrelevant example. Her answers (in turns 5 and 7) clearly established the need for a counterexample to satisfy the first condition of St35. As a consequence, Lizbeth became explicitly aware that irrelevant examples could not count as counterexamples since counterexamples should be taken from the set of analysis in the first place (her assumption dAL1[4]).

Introducing an irrelevant example to the discussion turned out to be a crucial step towards our goal of finding a description for the counterexamples that would disprove the given UAS. It allowed the teachers, and Lizbeth in particular, to become aware that a counterexample should necessarily satisfy the first condition in the statement.

[77] In Discussion 1.5 we had already established that St35 referred to *all people in the classroom* (see Appendix D1 for details on the task).

[78] Those two examples I made bold in the text of Episode 8 so they could be more easily located.

[79] If we think of the UAS *"All X are Y"*, an irrelevant example would be an example that does not satisfy *X*, the first condition of the UAS (for more details, see Chapter 3, Section I.2.2).

Episode 8

Turn	Who	What
1	Gessenia	*That it [a counterexample for St35] doesn't satisfy at least one condition.*
2	I	*What condition?*
3	Lizbeth	*To be minors, but then it would be that all are adults.*
4	I	*So, that it doesn't satisfy to be a minor? Ok then,* ***my father*** *is not a minor, does he count as a counterexample?*
5	Lizbeth	*Ah, but he is outside the classroom.*
6	I	*That's why I ask you all, what characteristics should this counterexample have in order to reject this statement?*
7	Lizbeth	*One of the people in this classroom, hmm, … who is an adult.*
8	I	*… An example that breaks, what part? This part (I pointed out the second condition of St35). It needs to satisfy that it is a person here in this classroom, but that it breaks the second part, which means that s/he should not be a minor, alright?*
9	Andrea	*Ahhhh!*
10	I	*It is not sufficient that it breaks that (second) part, but it should also satisfy the first part.*
11	Gessenia	*Specify one person present in this classroom, that is at least one person in this classroom.*
12	I	*What characteristics does it have? To be a person in this classroom?*
13	Gessenia	*A person who is present in the classroom.*
		…
14	I	*Let's suppose that there is* ***a 5-year-old boy*** *in here,* ***in this classroom****. Does this boy count as a counterexample for the statement given?*
15	Lizbeth	*That is not a counterexample. It is instead a confirmation, it is reinforcing the statement.*
16	I	*My question is, would it be a counterexample for this statement? A 5-year-old boy in this classroom?*
17	Lizbeth	*No.*
18	Gessenia	*It would not be a counterexample.*

The second challenging example (*"a 5-year-old boy in this classroom"*, turn 14) was a confirming example and it was introduced in response to Gessenia's apparent exclusive focus on the first condition of the statement (see turns 11 and 13). Lizbeth's refusal to accept the given example as a counterexample (*"that's not a counterexample. It is instead a confirmation, it is reinforcing the statement",* turn 15) is, again, evidence of her awareness that confirming examples did not qualify as counterexamples, which she displayed in a previous discussion (her assumption dAL1[2]). Clearly, Lizbeth could distinguish the different nature of, and the role played by, the given challenging examples in relation to St35.

Further signs of Lizbeth's awareness

Lizbeth was consistent with the use of her new assumption dAL1[4] that an irrelevant example could not be a counterexample throughout the intervention. For instance, during the recap for Discussion 1, at the beginning of the second day of the intervention, Lizbeth discarded the number 25 as a counterexample for the statement St41,

St41: *All numbers ending in the digit 3 are divisible by 4.*

Even though I tried to emphasize the fact that 25 was not divisible by 4 (i.e., that 25 did not satisfy the second condition of St41), Lizbeth explained that 25 could not be a counterexample for it, *"because the counterexamples must be numbers ending in the digit 3"*. This exhibited again Lizbeth's awareness that counterexamples should satisfy the first condition of the statement and therefore irrelevant examples did not qualify as counterexamples (her assumption dAL1[4]).

Lizbeth's assumption dAL1[5]: The relevance of the order in the statement

In Discussion 1.7 I put forward an irrelevant example that satisfied the second condition of the statement, in contrast to the irrelevant example I suggested in Discussion 1.6.1 (see Episode 8). As a result of the discussion, Lizbeth explicitly pointed out an important aspect when disproving UASs: the order in the statement. Lizbeth regarded the order in the statement as a relevant factor when determining whether an example was a counterexample or not.

During Discussion 1.7 the teachers were asked to determine whether the number 9 qualified as a counterexample for the (true) statement St37:

St37: *All numbers divisible by 6 are divisible by 3*.

The discussion was framed as a classroom episode where a teacher (María) made the claim that 9 (an irrelevant example) was a counterexample for St37. The number 9 had peculiar characteristics: it was an example that did not satisfy the first condition (9 is not divisible by 6), but it satisfied the second condition of the statement (9 is divisible by 3); in other words, it was a counterexample for the converse of St37, but not for St37.

Episode 9 focuses on Lizbeth's explanation for why the irrelevant example, number 9, was not a counterexample for St37, where she emphasized the order in the statement. Lizbeth identified that there was an order in the statement and this order played an important role when, for example, deciding whether an example qualified or not as a counterexample (turn 2). Another manifestation of Lizbeth's focus on the order in the statement was clear when she drew attention to the expectations for María and what she should have done first and next (*"First, she [María] has to find all numbers divisible by 6 and then see if they are also divisible by 3"*, turn 4). Lizbeth used this in particular to analyze the truth value of the statement; that is, if we generally think of the UAS *"All X are Y"*, Lizbeth planned first to focus on the elements in *X* and then analyze whether those elements belonged to *Y* or not. Presumably, she noticed that 9 satisfied the second condition and did not satisfy the first condition of the statement (i.e., it was a counterexample for the converse statement, but not for St37). This seems to have prompted Lizbeth to appeal to the order in the statement as a way to show her disagreement with the question of whether 9 was a counterexample for St37. It is likely that based on the analysis of the two challenging examples in Discussion 1.6.1 (see Episode 8 above) Lizbeth acknowledged the order in the statement as an important factor (her assumption dAL1[5]), in particular when determining whether an example was a counterexample or not.

Episode 9

Turn	Who	What
1	I	*In the episode María provided number 9 as she thinks that it is a counterexample. Is it a counterexample?*
2	Lizbeth	*Could it be that she [María] is not following the order in the statement?*
3	I	*What characteristics should that counterexample have?*
4	Lizbeth	*First, she [María] has to find all numbers divisible by 6 and then see if they are also divisible by 3. I found out that this is correct…*
5	I	*Alright, you are also analyzing the truth value of the given statement. Though, I'm not asking for that right now.*
6	Lizbeth	*Ah ok, the characteristics. Then the order, the order in the statement.*
7	I	*I mean, let's assume, if I were told that this statement was false, how should my counterexample be?*
8	Andrea	*It should be a number divisible by 6, but not divisible by 3. But there is no one [like that].*
9	I	*There is no one. Later we will see why, we will justify that in general. The thing here is that María thinks that 9 is a counterexample because she says, it is divisible by 3, but it is not divisible by 6. And that is not the statement she was given.*
10	Gessenia	*It is backwards.*
11	Andrea	*That means, it must satisfy the first [condition], but not the second [condition]*
12	I	*… If there was a counterexample, it would have to be a number divisible by 6, but not divisible by 3. María is confused and chose the other way around, that is, it satisfies the second condition and does not satisfy the first condition.*
13	Lizbeth	*That's why the order is very important.*

Further signs of Lizbeth's awareness

Lizbeth showed the use of her assumption dAL1[5] again during Discussion 2.3. The discussion revisited a classroom episode seen during Discussion 1.7, although the statement St37 was presented in its conditional form, that is, as St47,

> St47: *If a number is divisible by 6, then it is divisible by 3.*

Lizbeth's answer for whether 9 qualified or not as a counterexample for St47 was consistent with her assumption dAL1[5].

> Lizbeth: *I think no, because clearly if we have to follow the order here, first we have to find the numbers divisible by 6 so that just then, let's say, see if those numbers are also divisible by 3… 9 does not belong to the [set of] numbers divisible by 6.*

Lizbeth was consistent in using the order in the statement as a relevant factor to determine whether an example qualifies or not as a counterexample. It led her to conclude that 9 did not count as a counterexample for St47, because it was not divisible by 6 to begin with.

To sum up, Lizbeth managed to identify the characteristics of all possible counterexamples that would refute the discussed universal affirmative statements. Lizbeth could have reached a more concise definition of "counterexample" by combining three of her assumptions: the initial assumption that counterexamples should contradict something in a UAS (assumption dAL1[1]), plus the two aspects Lizbeth noticed about counterexamples (assumptions dAL1[2] and dAL1[4]), namely:

> (1) a counterexample should not simultaneously satisfy both conditions in the statement, and
>
> (2) a counterexample should satisfy the first condition of the statement.

Her assumptions were enough to infer a shorter and general description for all possible counterexamples, to wit, that *counterexamples should satisfy the first condition, but contradict the second condition of the statement*. Nevertheless, before Lizbeth accomplished that level of generality, Andrea did, presumably by relying on Lizbeth' observations, as I show next.

Figure 20 shows the development of Lizbeth's assumptions about the characterization of counterexamples to a UAS and the elements or aspects of the intervention that might have supported or triggered changes[80].

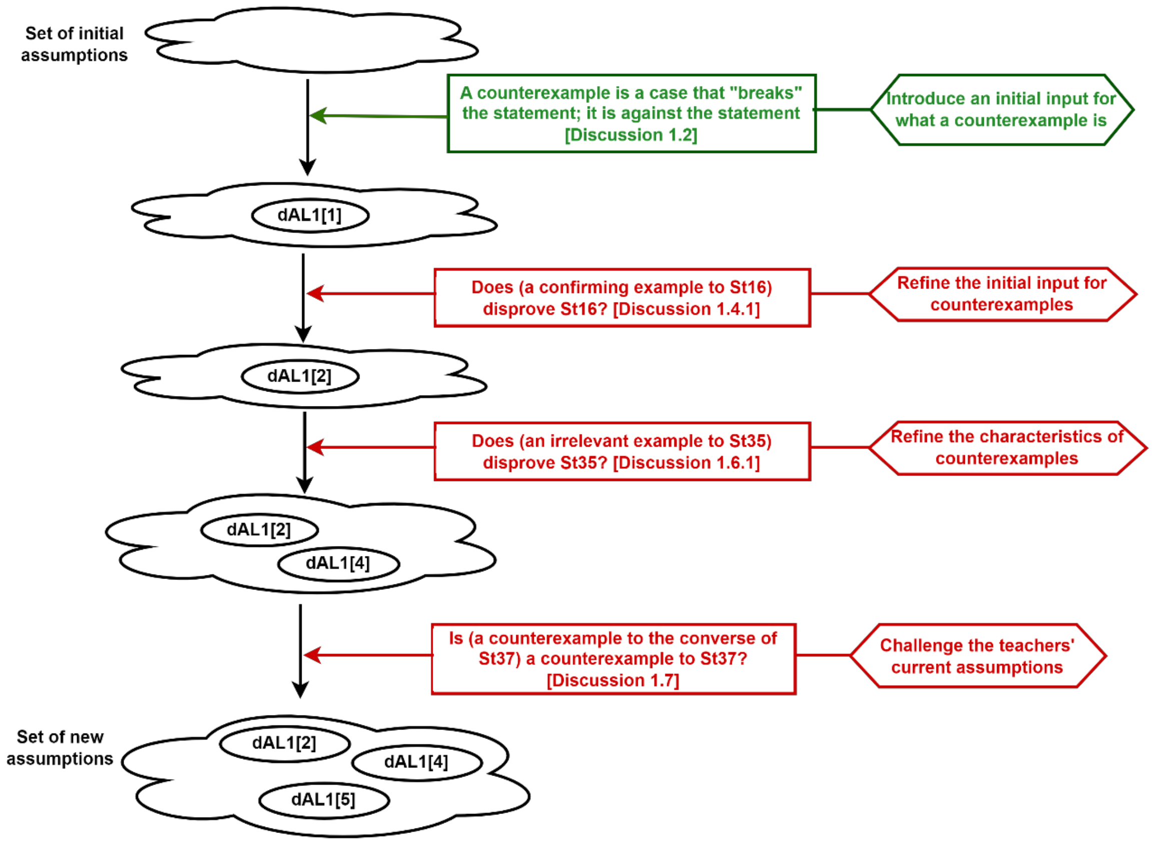


Figure 20. The development of Lizbeth's assumptions about the characterization of counterexamples to a UAS (dAL1[1]: A counterexample "breaks" the statement; dAL1[2]: It (a confirming example) cannot be a counterexample to a UAS (St16) because it supports the statement; dAL1[4]: It (an irrelevant example) cannot be a counterexample to a UAS (St35) because it does not satisfy the first condition of the statement; dAL1[5]: The order in the statement is important)

[80] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

2.2.2. Andrea's assumptions

Andrea was the only teacher who exhibited an initial assumption for what a "counterexample" is. Before the intervention Andrea explicitly said that she used the term "counterexample" in her teaching of mathematics; however, as I show below, her usage of the term differed from what is typically understood in a mathematical context.

Table 10 includes Andrea's assumptions related to the characteristics of counterexamples to a UAS.

Table 10. Andrea's assumptions about the characteristics of counterexamples to a UAS

Assumption	**Type of assumption (when first observed)**
bAA4: A counterexample is a conflicting example or a conflicting situation	Initial assumption (before the intervention)
dAA1[5]: A counterexample should "break" the statement	New assumption (during the intervention)
dAA1[9]: A counterexample must satisfy the first condition, but contradict the second condition of the universal statement	New assumption (during the intervention)
dAA1[9*]: The counterexample must be always taken from the antecedent of the universal statement	New assumption (during the intervention)
dAA1[9**]: The counterexample must satisfy the antecedent and not the consequent of the universal statement	New assumption (during the intervention)
aAAm1: A verbalized semi-general[81] counterexample is valid to disprove a US	New assumption (after the intervention)

Illustrations of Andrea's use of her initial assumption as well as the development of her subsequent assumptions about the description for counterexamples and cases where she applies her new insights are given next.

Andrea's initial assumption

Andrea's personal initial use of the term "counterexample" (bAA4)

During the exploratory interview, before the intervention, the teachers were asked how they normally verified their students' understanding of the mathematics lessons they taught. Andrea was the only teacher who explicitly referred to "counterexamples" when answering this question.

> Andrea: … *I also ask them, why do you add? Or, why do you put these together? I always ask them why and I provide examples and counterexamples to check whether they are understanding what they are doing, because sometimes they only link the numbers and say, ah ok, the answer is one. But I tell them, no, what about if we subtract? Then, as I give them this challenge, providing counterexamples to them, so they are sure that yes, indeed, this should be added.*

The way Andrea used the term *counterexample* in her teaching of mathematics suggested the idea a "conflicting example" or a "conflicting situation". Her use of "counterexample" is related to a type of uncertainty that Zaslavsky (2005) discussed and that is entailed in certain mathematical tasks, namely, *competing claims*[82]. In Andrea's case, she introduced a contradicting operation as a way to challenge her students' certainty about their

[81] Andrea did not use the term "semi-general counterexamples", but the case she referred to qualified as one.
[82] For details, see Chapter 2, Section III.2.

mathematical choices. This means that she suggested an operation she knew beforehand was not appropriate to tackle a given problem. The unsuitable operation would be what she called a "counterexample". In this sense, for Andrea a "counterexample" was an example (of a number, an operation, etc.) somehow against her students' answers or solutions to a problem analyzed in class.

From this view, Andrea's initial use of "counterexample" was not in accord with its use in mathematics, which is the sufficient justification to disprove a universal statement. Evidence of this was Andrea's lack of awareness that the example given by Lizbeth during Discussion 1.0 to disprove a UAS was indeed a counterexample (see Section I.2.1.2 above). This means that Andrea did not use the term "counterexample" in the context of mathematical statements and justifications. Instead, her usage aimed at challenging her students' choices for the operations they used to solve mathematical tasks.

Andrea's general characterization of counterexamples to a UAS (dAA1[9])

Andrea's genuinely surprised reaction (turn 9 in Episode 8, above) to my summary of Lizbeth's answer is evidence that she was not aware yet of the characterization for all possible counterexamples to St35.

> St35: *The people present in this classroom are minors*

This discussion constituted an important moment for Andrea, who seemed to have internalized this as a new general assumption of her own, which she explicitly articulated later, during Discussion 1.7.

In Discussion 1.7 Andrea expressly established a general assumption (her assumption dAA1[9]) that summarized a characterization for all possible counterexamples that would disprove a given (false) UAS (*"it must satisfy the first [condition], but not the second [condition]"*, turn 11 in Episode 9, above). This was a generalization of the characteristics previously identified by Lizbeth for concrete statements in previous discussions combined with a generalization of the characteristics Andrea herself identified for the (impossible) counterexamples to St37 *"All numbers divisible by 6 are divisible by 3"*.

> Andrea: *It should be a number divisible by 6, but not divisible by 3* (turn 8 in Episode 9).

Notice that Andrea was aware that there was no counterexample to St37 ("*But there is no one [like that]*", turn 8 in Episode 9); however, Andrea was able to provide a characterization for those nonexistent counterexamples.

Two later discussions show Andrea's adaptation of the terminology introduced for conditional statements to her general assumption dAA1[9]. During Discussion 2.3 Andrea reformulated part of her assumption as *"The counterexample always must be [taken] from the antecedent"*. Her "updated" formulation of dAA1[9] replaced the *"must satisfy the first condition"* part of her complete assumption with *"must be taken from the antecedent"*. To this I called her assumption dAA1[9*] in Table 10. Later, in Discussion 7.1, Andrea provided a more complete description for counterexamples according to the terminology for conditional statements: *"the counterexample must satisfy the antecedent and not the consequent"* (dAA1[9**]). Andrea decided to modify the formulation of her assumption dAA1[9] according to the new concept (conditional statements) and its related-terminology; however, she kept the main core of the assumption. This is evidence of her consistent use of the general description for counterexamples she had concluded (her assumption dAA1[9]).

This suggests that the discussion about whether irrelevant and confirming examples disproved UAS successfully triggered the formulation of a general description for all im/possible counterexamples to UASs. In the particular context of my intervention, this sort of analyses supported the teachers' teamwork towards that goal. Lizbeth's rejection of irrelevant and confirming cases for specific statements, plus her explanations for why she discarded them as counterexamples, allowed Andrea later to articulate a general description.

Andrea's application of her assumption dAA1[9]

There are two interesting contexts in which Andrea exhibited her application of her assumption dAA1[9], that a counterexample should satisfy the first condition, but not the second condition of a universal statement. The first was her search for counterexamples when evaluating a (hard-to-disprove) universal *negative* statement. The second was her acceptance of a *non-minimal* disproof through which she introduced a new concept: "verbalized counterexamples". Both cases revealed her flexible current broad understanding of counterexamples.

Andrea's looking-for-counterexamples approach

During Discussion 3 the teachers were asked to solve a task that consisted in observing a pattern, formulating a conjecture, determining whether their conjecture was true or not, and explaining why[83]. The teachers conjectured that:

"For every natural number n *(different from zero), when substituted in the formula* $1 + 1141n^2$*, the result is not a perfect square number"*

While, Gessenia and Lizbeth seemed very confident that the conjecture was true, based on a number of examples they verified, Andrea tried to find a counterexample (for details, see Section I.3 below). Her *looking-for-a-counterexample* approach consisted in equating the given expression to a square number, or as she showed in her notes:

$$1 + 1141n^2 = x^2$$

Her formulation was already evidence of Andrea's use of her assumption dAA1[9] since she tried to look for a natural number *n* such that the expression $1 + 1141n^2$ was a perfect square number. This revealed her search for a counterexample, that is an example that satisfied the first condition (*n* is a natural number different from zero) and did not satisfy the second condition of the statement ($1 + 1141n^2$ is not a perfect square number). An interesting aspect of Andrea's reasoning is that she extended her understanding about the characterization of all possible counterexamples to a UAS to the case of a conjecture that was a universal *negative* statement (UNS). At this point of the intervention we had not yet discussed this type of universal statement. I return to this point later in Section I.3.

[83] For details, see "Task" in Appendix D3.

Andrea's acceptance of a non-minimal justification: The case of "verbalized counterexamples"

During our Meeting #9 the teachers were asked to analyze arguments for different statements. Here my focus is on the statement St45.

> St45: *If a distribution is whole, fair and maximal, then there are zero objects left*

One of the tasks consisted in analyzing Argument T2 for St45.

> Argument T2: *This is not true, because if you have less objects than people, it is indeed possible that we have [one or more] objects left.*

Observe that Argument T2 is a general description for a class made of some of the possible counterexamples for St45 (it is a *semi-general counterexample* in Peled & Zaslavsky's [1997] terms). It is also a non-minimal justification as produces more than one counterexample to disprove St45. Andrea's response addressed both the mode of argumentation and the form of expression of the argument.

> Andrea: *I believe this is a verbalized counterexample (she smiled). For me this is valid. They have not included numbers, but what they have stated is a counterexample for what has been claimed [in the statement]. It has just been verbalized.*

Andrea claimed that the argument qualified as a *valid "verbalized counterexample"*. She called the argument "verbalized" because of the verbal mode of argument representation used (A. J. Stylianides, 2007); that is, no numerical or algebraic expressions were included, but only words. Andrea recognized that it was a *counterexample* because she was aware that it was a description that may generate counterexamples since instances of the argument satisfied the conditions from her assumption dAA1[9], namely, a FWM-distribution with a number of objects left different from zero. These conditions were actually part of a property the teachers proved before: *in a FWM-distribution where the number of objects is smaller than the number of people, the number of objects received by each person is zero and the number of objects left is the same as the number of objects that were given to distribute*. In fact, with these conditions, infinite counterexamples could be given. Considering all this, Andrea was aware that the argument involved a suitable mode of argumentation (A. J. Stylianides, 2007) to disprove St45 (*"For me this is valid"*).

The prior literature reports cases of individuals who reject non-minimal correct disproofs for universal statements (Tabach et al., 2010b; Tsamir et al., 2009)[84]. In this case, even though Andrea was aware of the "minimal justification" required to refute a UAS or UCS (her assumption dAA1[3] in Section 2.1.2 above), Andrea's understanding was flexible enough to not reject Argument T2 because "it did not follow the needed framework" or because it involved "overdoing" the sufficient justification, as some teachers might think (e.g., see Tsamir et al., 2009). Instead, Andrea focused on the mode of argumentation since she paid closer attention to the conditions included in Argument T2 to decide whether it qualified or not as a counterexample for St45. This suggests that Andrea prioritized the mode of argumentation over the form of expression. She was aware of the characterization for all possible counterexamples (her assumption dAA1[9]), in particular for St45, and as such she could discern whether the argument entailed such a description or not. Andrea's criterion to judge Argument T2 as a disproof for St45 was focused on

[84] For details, see Chapter 2, Section I.5.2.

the mode of argumentation and it overrode the minimality of the argument, of which she was also aware.

It is interesting that even though no cases of "verbalized counterexamples" or non-minimal justifications were put forward in previous discussions during or after the intervention, Andrea did not carelessly disregard Argument T2 without considering first the mode of argumentation involved in it. Furthermore, Andrea's explanation suggested that she was confident in her acceptance of the argument. This is indicated by her use of the adjective "valid", which she used to state that Argument T2 was correct and which showed Andrea's certainty about it to disprove St45. With her concept "verbalized counterexample" Andrea showed two of her current insights about counterexamples: 1) a counterexample should satisfy the first condition of the statement and contradict the second condition (her assumption dAA1[9]), as Argument T2 verified those conditions, and 2) a counterexample is sufficient evidence to disprove a US (her assumption dAA1[3] in Section I.2.1.2), but there might be other types of counterexamples, like "verbal counterexamples", that are also valid to disprove a US (her new assumption aAAm1).

Andrea's teaching as evidence of her own understanding

During her teaching of Session 10 Andrea's class was engaged in solving a task that involved the (false) universal statement St143,

> St143: *All natural numbers are divisible by 4*.

Andrea asked her students to determine whether the division of 0 by 4 (a confirming example for St143) would be a counterexample for St143. They had already solved the division on the whiteboard in front of the class.

> Andrea: *Then this [St143] is false, then the answer is "no", and which one is your counterexample? Is this one (she points to the division of 0 by 4 on the whiteboard) your counterexample?*
>
> Students: *No*
>
> Andrea: *No, because that supports this, this says yes-*
>
> Student G: *That is an example!*
>
> Andrea: *That is an example, but the other one is a counterexample, that means, that is sufficient to justify the statement [is false].*

Andrea pointed out that the division that she suggested was not a counterexample because it supported the statement. Student G even emphasized that it was instead an example, in the sense of a confirming example. Andrea's instruction consisted in making the distinction that a confirming example could not be a counterexample, which is consistent with her assumption dAA1[9] that a counterexample must satisfy the first condition in the statement and contradict the second condition. It also shows that the approach taken during the intervention for teachers seemed to help Andrea to reject confirming examples as counterexamples that could refute UASs.

Likewise, during her teaching of Session 13 Andrea had not only the chance to emphasize that a confirming example is not a counterexample, but also to highlight the condition needing to be denied. Andrea's class had identified that statement St149 is false and discussed the justification that disproved it.

> St149: *All numbers divisible by 3 are bigger than 3*

Andrea concluded the whole-class discussion by writing down on the whiteboard in front of the class: *"No, because 0 and 3 are divisible by 3, but they are not bigger than 3"*. However, she noticed that one student suggested a confirming example (the number 6), which pushed her to engage her class in a further discussion for clarification.

> Andrea: *The number 6 is bigger than 3, so it does satisfy [the second condition of St149]. Remember, here it is stated that (Andrea reads St149 again) all numbers divisible by 3 are bigger than 3. The number 6 is bigger than 3, it does satisfy what is claimed in the statement, but if you say that this is not true, you have to tell me which ones do not satisfy.*

Andrea expected that her students provided an example that did not satisfy the second condition in St149 (*"The number 6 is bigger than 3, ... but if you say that this is not true, you have to tell me which ones do not satisfy"*). Her feedback pointed to the condition that should not be satisfied; that is, the condition *"is bigger than 3"* (*"The number 6 is bigger than 3, it does satisfy what is claimed in the statement, but if you say that this is not true, you have to tell me which ones do not satisfy"*). This is clearly consistent with her assumption that a counterexample should not satisfy the second condition in the statement (part of her assumption dAA1[9]).

Irrelevant examples cannot be counterexamples

The emphasis on the first condition of a statement was clear in Andrea's teaching of her Session 14. The students were supposed to solve Activity 10 individually first[85]. This activity consisted of two main tasks. The second task, which is the focus here, was made up of two items. Both items referred to the same imaginary situation: A magic bag that contained all numbers divisible by 4. The students had to imagine themselves taking away numbers from the bag, one by one. The first item asked the students whether *some of those numbers were divisible by 8*, whereas the second item asked the students whether *all of those numbers were divisible by 8*. For both items, they were expected to choose a "yes" or "no" answer and explain their choice. Episode 10 (below) is centered on the whole-class discussion about the students' answers to the second item, which entails statement St151,

> St151: *All numbers divisible by 4 are divisible by 8.*

The students' answers were divided. Some considered that indeed, all those numbers were divisible by 8, and some students chose the "no" answer. Before the discussion in Episode 10 one student had already suggested the number 44 (a counterexample) to the class in order to refute St151.

Andrea's teaching can be split into three main parts. First, Andrea wanted to find out whether the students were aware of the conditions the example to be provided should or should not satisfy (turns 1 and 3). Second, as she noticed that at least one student could not identify that the evidence should satisfy the first condition in the statement (Student X claimed that both conditions should not be satisfied, see turn 2), she drew the students' attention to the statement itself and its respective interpretation in order to identify the characteristics of the example she expected (turns 3 and 5). Third, Andrea provided an irrelevant example, 5, so that her students could reject it as disproving that *"All the numbers removed from the bag are divisible by 8"* (turn 11).

[85] Appendix A10 includes Activity 10 as it was presented to the students.

Episode 10

Turn	Who	What
1	Andrea	*But was I supposed to show a number that is only not divisible by 8?*
2	Student X	*It should not be divisible by 4, neither should it be divisible by 8.*
3	Andrea	*Kids, here it says that you will take away numbers from the bag. If I am going to take a number from the bag, what should it satisfy?*
4	Student Y	*It should be divisible by 4, but it should not be divisible by 8.*
5	Andrea	*Exactly! That is, here you are asked about all the numbers you take away from the bag, it means, AAAAALL the numbers that are in here, in the bag, that are divisible by 4. Are they divisible by 8?*
6	Students	*No.*
7	Andrea	*No, because here we have 44. Miss, 44 is divisible by 4, it is indeed in the bag, but it is not divisible by 8. What am I automatically doing with this example?*
8	Students	*Breaking.*
9	Andrea	*What am I breaking?*
10	Students	*The statement.*
11	Andrea	*Could you prove this is false by showing* ***number 5****?*
12	Students	*No!*
13	Andrea	*No, because, was 5 in the bag?*
14	Student X	*Miss, we could have chosen number 4 too.*

Notice that the irrelevant example that Andrea suggested did not satisfy either the first (it was not a number in the bag) or the second condition of the statement (it was not divisible by 8)[86].

Based on her intervention, Andrea intended not only to make clear that an irrelevant example could not disprove the statement under discussion, but also that her students understood why. With that purpose she made use of the logical interpretation of the statement (*"That is, here you are asked about all the numbers you take away from the bag, it means, AAAAALL the numbers that are in here, in the bag, that are divisible by 4. Are they divisible by 8?"*, turn 5).

Figure 21 shows the development of Andrea's assumptions about the characterization of counterexamples to a UAS and the elements or aspects of the intervention that might have supported or triggered changes[87].

[86] In fact, in this case it was impossible for her to provide an example that did not satisfy the first condition, but satisfied the second condition in the statement given that the second condition implied the first.

[87] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

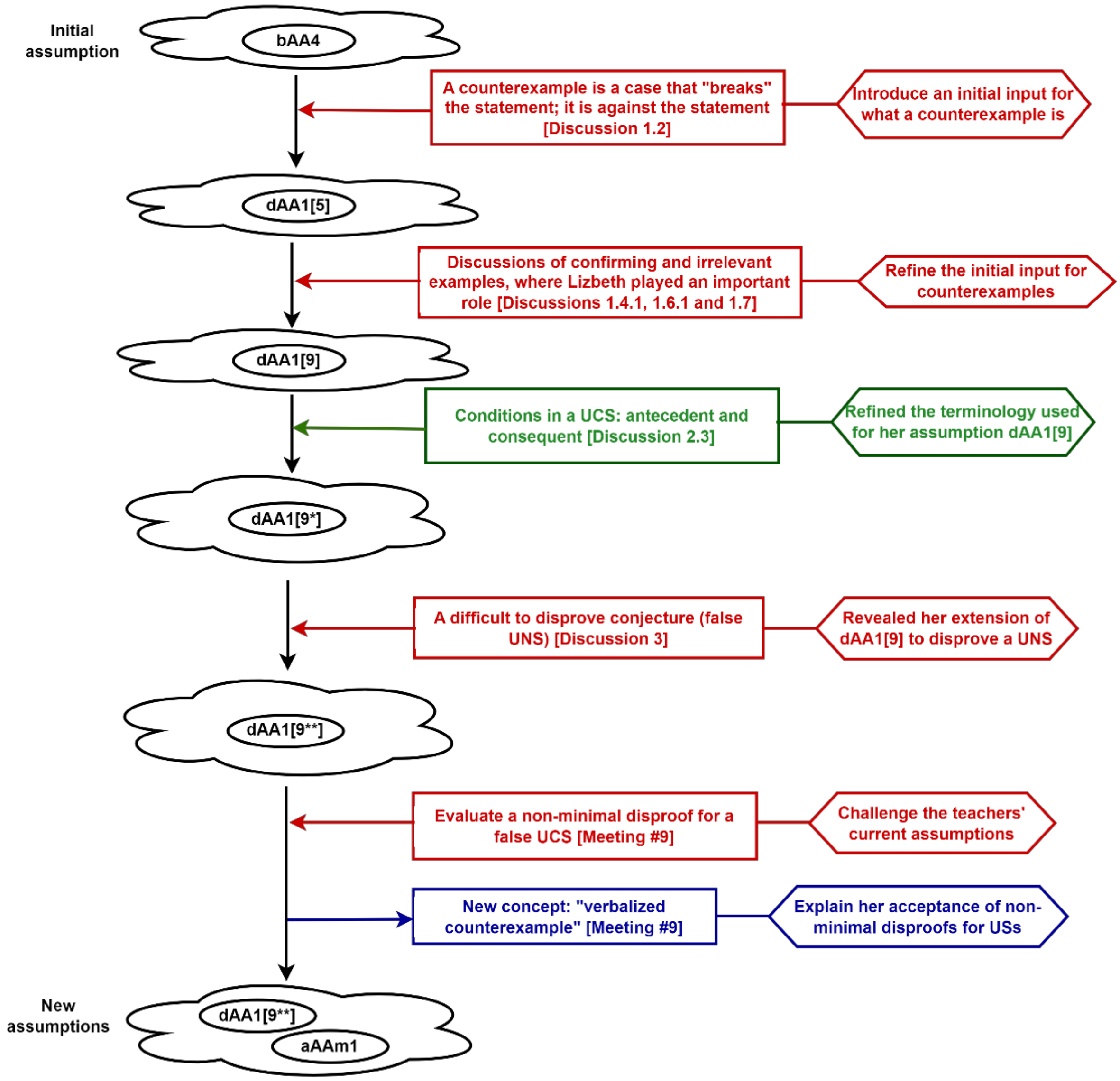


Figure 21. The development of Andrea's assumptions about the characterization of counterexamples to a UAS (bAA4: A counterexample is a conflicting example or a conflicting situation; dAA1[5]: A counterexample should "break" the statement; dAA1[9]: A counterexample must satisfy the first condition, but contradict the second condition of the universal statement; dAA1[9]: The counterexample must be always taken from the antecedent of the universal statement; dAA1[9**]: The counterexample must satisfy the antecedent and not the consequent of the universal statement; aAAm1: A verbalized semi-general counterexample is valid to disprove a US)*

2.2.3. Gessenia's assumptions

Gessenia began the intervention for teachers with no initial assumptions about what a counterexample was. This means that any assumption she revealed was intervention-based.

Table 11 includes the assumptions that Gessenia developed during and after the intervention about the characteristics of counterexamples to UASs.

Table 11. Gessenia's assumptions about the characteristics of counterexamples to a UAS

Assumption	**Type of assumption (when first observed)**
dAG1[3]: A counterexample should "break" the statement.	New assumption (during the intervention)
dAG1[4]: Counterexamples should not satisfy at least one condition in the statement.	New assumption (during the intervention)
dAG1[5]: Counterexamples should contradict the statement.	New assumption (during the intervention)
aAGt2: A counterexample should verify the first condition, but contradict the second condition of the UAS.	New assumption (after the intervention)

Her assumption dAG1[3] is the implicit assumption I presume all the teachers began the intervention with, since it is based on the initial input I provided during Discussion 1.2. In Gessenia's case, she explicitly used this initial input in different forms, in particular as her assumption dAG1[5]. Her assumption dAG1[4] revealed her imprecision when determining the exact condition of a statement that counterexamples should deny. Gessenia used this assumption throughout the intervention, in spite of the discussions we had where Andrea and Lizbeth pointed out important factors such as the order in the statement (see Section I.2.2.1 above). Nonetheless, this is also evidence that Gessenia did not adopt others' observations, like Andrea's assumption that a counterexample should satisfy the first but not the second condition of the statement, without first grasping them.

Gessenia's imprecise first assumptions (dAG1[4] and dAG1[5])

Counterexamples should not satisfy at least one condition in the statement (dAG1[4])

During Discussion 1.6.1 the teachers were expressly asked for the characteristics of the counterexamples that would refute the statement St35.

St35: *The people present in this classroom are minors*

Gessenia's answer (*"That it doesn't satisfy at least one condition"*, turn 1 in Episode 8 above) revealed one of her current assumptions. In her answer, Gessenia did not precisely identify a condition that a possible counterexample should contradict. Instead, her answer was imprecise, which was directly influenced by the way I introduced the term "counterexample" during Discussion 1.2 ("A counterexample is a case that "breaks" the statement; it is against the statement").

Discussion 1.7 exhibited Gessenia's application of her assumption dAG1[4] in a different context. Episode 11 (below) shows Gessenia's affirmative answer to the question of whether number 9 qualified as a counterexample for the statement St37,

St37: *All numbers divisible by 6 are divisible by 3.*

Episode 11

Turn	Who	What
1	Gessenia	*Yes… Because we have said that in order to be a counterexample it is sufficient that [---].*
2	I	*What characteristics?*
3	Gessenia	*That it does not satisfy the condition.*
4	I	*Again, be more precise.*
5	Gessenia	*One can see here that 9 is not divisible by 6… Here where it says the numbers divisible-, we are talking about all the numbers divisible by 6, (they) are divisible by 3. But we see that not all numbers divisible by 6 are divisible by 3.*
6	I	*What characteristics should the number I must provide have in order to guarantee that it is a counterexample that is going to reject that statement?*
7	Gessenia	*That it is not a multiple of 6… That is not divisible by 6. It says, all numbers divisible by 6 are divisible by 3. It means, that number must be divisible by the two numbers, by 6 and by 3.*
8	I	*In order to be a counterexample?*
9	Gessenia	*No, in order to be a counterexample, it should not be divisible by at least one of them, by 6 or by 3.*

Gessenia's explained the rationale behind her answer as *"in order to be a counterexample it should not be divisible by at least one of them, by 6 or by 3"* (turn 9). Gessenia was aware that the first condition of St37 was not satisfied was, (*"One can see here that 9 is not divisible by 6"*, turn 5). Thus, Gessenia decided that 9 was a counterexample for St37 since 9 did not satisfy at least that condition (it was not divisible by 6).

Furthermore, her assertion in turn 7 implies that Gessenia understood St37 as the conjunction *the numbers are divisible by 6 and divisible by 3* (*"that number must be divisible by the two numbers, by 6 and by 3"*, turn 7). Hence, her claim that the statement would be false if at least one of the conditions was not satisfied (turn 9) seems reasonable as a conjunction is false if at least one of the simple clauses that make the conjunction is false. This would also explain why Gessenia's assumption was very strong and resistant to change despite the previous discussions where Lizbeth had already begun to highlight important aspects about counterexamples (see Section I.2.2.1 above). Gessenia understood that St37 is a universal statement because she explicitly referred to the universal quantifier "all" (*"we are talking about all the numbers divisible by 6"*, turn 5); however, she interpreted it as a conjunction where the order of the simple clauses was irrelevant. Her initial interpretation of St37 as a conjunction seems to have suggested her that 9 was a counterexample for St37, given that a counterexample *"should not be divisible by at least one of them, by 6 or by 3"*, without paying attention to the order of the clauses.

Gessenia's rationale was consistent with her assumption that a conditional statement and its converse assert the same thing[88]. If Gessenia interpreted a UAS as a conjunction, where the order of the clauses is not important, then this could also explain why she made no distinction between a UAS and its converse. In this case, Gessenia actually provided a counterexample for the converse of St37 and not for St37. Other researchers have reported cases where individuals also accepted a counterexample of the converse as if it were a

[88] For details on Gessenia's assumption that a universal conditional statement and its converse asserted the same, see Section I.1.1.2 above.

counterexample for the original statement (e.g., Zaslavsky & Ron, 1998), even though they were aware that one counterexample was sufficient to refute the statement.

Episode 9 (above) shows the continuation of Episode 11. In Episode 9 Lizbeth pointed out the order in the statement as a factor that was not considered when 9 was suggested as a counterexample (turns 2, 4 and 6). Andrea provided the characteristics for the counterexample that should have been provided (turn 8). Presumably, those, and the feedback I gave afterwards, pushed Gessenia to come to realize that 9 was indeed a counterexample, but for the "backward" statement (turn 10). However, whether it was really relevant for her or not, and whether she integrated the factor *order in the statement* to her considerations related to counterexamples was uncertain since she did not provide any further comment about it.

Counterexamples should contradict the statement (dAG1[5])

The following day Gessenia showed new evidence of imprecisions in her assumption about a characterization for counterexamples. During a Recap of Discussion 1 the teachers were asked to provide one false universal statement each. At this point, the three teachers were aware that a counterexample would be sufficient to prove that the statement was false. Attention was then drawn to Gessenia's statement St40,

St40: *All even numbers end in an odd-digit*.

The teachers were asked to identify the characteristics of the counterexamples that would disprove St40. Gessenia answered *"one that contradicts the statement"* (I called it her assumption dAG1[5]), which showed again her imprecision. As in the case of her first assumption dAG1[4], that a counterexample should contradict at least one condition in the statement, here she did not specify what condition a counterexample should contradict. On the contrary, her answer was still broad and exhibited that Gessenia did not pay attention to the same aspects Andrea and Lizbeth did throughout Discussion 1.

Gessenia's participation in Discussion 8 revealed again her lack of focus on the structure of a statement in order to suitably identify a counterexample for it. At this point the discussion was focused on evaluating the truth value for statement St119.

St119: *If A is divisible by B, then B is divisible by A*

Andrea provided the case *4 divided by 2* (i.e., $A=4$ and $B=2$) as a counterexample that "broke" (disproved) the statement. In order to support Andrea's answer, Gessenia tried to provide other examples that, according to her, "contradicted" the statement. Among them, she gave the irrelevant example *3 divided by 5* (i.e., $A=3$ and $B=5$) and regarded it as a counterexample. Gessenia might have provided this example for two reasons: (a) Gessenia used again her broad assumption that a counterexample should contradict "something" in the statement (in this case both conditions: 3 was not divisible by 5, neither was 5 divisible by 3), or (b) she did not understand the statement. The former explanation suggests that Gessenia had not been influenced by the previous discussions we had, in which Lizbeth and Andrea put emphasis on the order in the statement as an important aspect when identifying the characteristics of counterexamples; that is: counterexamples should satisfy the first condition of the statement, but contradict its second condition (see Sections I.2.2.1 and I.2.2.2 above). In the latter the structure of St119 was possibly difficult to distinguish. She needed to realize first that the set of analysis was made of pairs of elements A and B $(A;B)$ that satisfied the condition that A was divisible by B. This might not have been that evident to Gessenia. In order to check this, I drew the teachers' attention to the identification of the set of analysis for St119 and an equivalent statement for it, where its universal nature was more transparent. Gessenia provided examples that exhibited her lack of understanding of the logical interpretation for St119 (e.g., single numbers like 10).

I also contrasted St119 with the statement *"If A is divisible by B, then there exist cases in which B is divisible by A"*, which was also part of Discussion 8. Attention was particularly drawn to the logical interpretation of the statements. Altogether, these clarifications might have supported Gessenia's emerging understanding. Much later, during her teaching she showed a more precise insight into exactly what conditions counterexamples of a universal statement should and should not satisfy.

Traces of emerging understanding: Gessenia's Teaching

Evidence of Gessenia's evolving understanding of the description for all possible counterexamples was observable only after the intervention for teachers, where she also used a similar approach to the one I used during the intervention for teachers: introduce confirming and irrelevant examples and request her students to evaluate whether they qualified or not as counterexamples.

Gessenia's new assumption aAGt2: Counterexamples must verify the first, but contradict the second condition of the statement

During her teaching of Session 7 Gessenia introduced the term "counterexample". Gessenia's class discussed the truth value and justification for the statement St45.

> St45: *If a distribution is fair, whole and maximal (FWM), then there are zero objects left*

Gessenia invited her students to discuss one group's answer that she chose because it was a correct answer (the students recognized that the given UCS was false) with a sufficient justification (it included a counterexample).

> Gessenia: *Is that an example? ... Are they doing what the statement states? Are they giving a division with remainder zero? Are they providing a division with remainder zero? They are doing the opposite… they are using a counterexample… they say NO, and in order to prove that this is not true, they do … the opposite… They say, zero left? NO. I will show you that there is not always zero left. Here it is [she points to the division "17 divided by 3" that is in front of the class on the whiteboard], two are left. It is a FWM-distribution and there are two left, remainder two.*

The way Gessenia introduced the term "counterexample' was based on a contrast she emphasized with confirming examples. First, she focused on the characteristics of a confirming example (*"Is that an example? ... Are they doing what the statement states? Are they giving a division with remainder zero? Are they providing a division with remainder zero?*); then she highlighted that the group's answer went into a different direction, which allowed her to introduce "counterexamples" as the examples that "opposed" the statement (*"They are doing the opposite… they are using a counterexample… they say NO, and in order to prove that this is not true, they do… the opposite"*).

Gessenia's introduction of "counterexamples" as opposing the statement might seem broad; however, the fact that Gessenia explicitly listed the characteristics of the "opposing" examples revealed her awareness of the characteristics that the possible counterexamples for St45 should have. Gessenia emphasized that the evidence consisted of an example that contradicted the second condition of the statement (*"I will show you that there is not always zero left. Here it is, two are left."*) and satisfied the first condition (*"It is a FWM-distribution"*). This already displayed Gessenia's awareness of a more precise description for counterexamples in comparison with her broad assumption that *counterexamples should contradict the statement* during the intervention for teachers (her assumption dAG1[5]). Moreover, her new assumption aAGt2 was aligned with Andrea's general description of all possible counterexamples of a US (a counterexample should satisfy the first condition of the statement and contradict the second condition, see Section I.2.2.2 above).

An important illustration of her emergent understanding was evident during Meeting #11 where Gessenia identified the characteristics of *hypothetical*[89] counterexamples for the imaginary statement St140.

St140: *All Peruvians are BLUMEN*

Gessenia's first attempt included an imprecise claim (*"at least an example that contradicts the statement"*); however, when pushed to be more specific, she added: *"at least one Peruvian that is not BLUMEN"*. This discussion was particularly interesting as it involved the case of an imaginary statement, for which its truth value was impossible to determine and which might have become a challenge for the teachers as they could not rely on the context and/or content. Through this discussion Gessenia revealed her awareness of the characteristics that counterexamples should have for UASs in general (her assumption aAGt2) through the characterization of counterexamples to the imaginary UAS St140. Gessenia managed to provide a description of the hypothetical counterexamples for St140, even though she needed an explicit request to be more specific with her characterization.

In addition, Gessenia applied her new assumption to a new statement during her teaching of Session 12. In her lesson, Gessenia asked her students to solve two tasks[90]. The first task involved completing statements of the form "16 is divisible by ____" with specific numbers so that the statement was true. The class called such statements "true sentences" and the students were requested to find all possible true sentences for 16. The second task included a classroom episode in which a boy called Julito explored other true sentences for numbers different from 16. He found the number of all possible "true sentences" for numbers 3, 6, and 24. Julito noticed that as the number increased (3, 6, 16, 24), the number of "true sentences" increased as well (two, four, five and eight "true sentences", respectively). Julito came up with conjecture St70,

St70: *Every time we use a bigger number, we will always get more true sentences than for the previous numbers.*

Gessenia's students were expected to identify whether Julito's conjecture was true or false and justify their answer. Several students considered that Julito's conjecture was true and some of them regarded conjecture St70 as a mathematical truth. Gessenia asked the students questions to find out and then refine what they understood by *conjecture* and *mathematical truth*. Even though some students had already noticed that the conjecture

[89] Counterexamples to St140 can only be *hypothetical* given that St140 is an imaginary statement and as such its truth value cannot be determined. Hence, one can only talk about counterexamples to St140 hypothetically. For more details on types of counterexamples, see the introduction of Section I.2, above.
[90] For details on the activity as it was presented to the students, see Appendix A12.

was not true, Gessenia requested the whole class to explore new numbers. One student came up with the number 25 as an example that "broke" Julito's conjecture St70. Gessenia used this example for the whole-class discussion:

> Gessenia: *He [Julito] says that always and we have seen that not always. An example, an example that breaks the claim he made is enough… Here it says, Julito says, every time we use a bigger number, always, he does not make any exceptions, he does not say this does not or the other one does not. He says, bigger numbers, always. He does not say, in these cases it does verify, in these cases it does not, he says we will ALWAYS get more true sentences and we have seen that this is not true, number 25 is bigger than 24 and it [25] does not have more true sentences than 24. The number 24 has eight true sentences and the number 25 has three true sentences. Therefore, what Julito says is not true.*

The first part of Gessenia's feedback displays her awareness about the sufficiency of one counterexample to disprove a UAS[91]; whereas the second part of her feedback shows that Gessenia was aware that a counterexample for a statement should satisfy the first condition of the statement (*"number 25 is bigger than 24"*) and contradict its second condition (*"it [25] does not have more true sentences than 24. Number 24 has eight true sentences and number 25 has three true sentences"*). In this context, 25 was a counterexample that disproved St70 (*"Therefore, what Julito says is not true"*).

An irrelevant example is not a counterexample

In her teaching of Session 14 Gessenia put forward an irrelevant example to her class so that her students determined whether the example disproved or not statement St151[92].

> St151: *All numbers divisible by 4 are divisible by 8*

Once the students identified that the statement was false, Gessenia used the irrelevant example 7 to draw her students' attention to the requirement that a counterexample is taken from the set of analysis. She then provided actual counterexamples to point out the characteristics that counterexamples for St151 should have.

> Gessenia: *Is it okay if you chose the number 7 to conclude that this [St151] is false? ... I ask, is 7 in the bag? ... We are working with all the numbers divisible by 4… in item "b" it was "all", but what does it refer to when it says "all"? ... to this group (she points to the bag she drew on the whiteboard where she tried to represent that all the numbers divisible by 4 were inside), here it says all these, all the numbers divisible by 4; it says, all the numbers I take away from the bag verify that they will be divisible by 8? ... No, because 4 is in the bag and 4 is not divisible by 8, 20 is in the bag and 20 is not divisible by 8, 12 is in the bag and 12 is not divisible by 8…*

Gessenia also focused on the logical interpretation of the statement to support her students' sense making of why irrelevant examples did not disprove an *"all-statement"* (*"in item 'b' it was 'all', but what does it refer to when it says 'all'?"*). She highlighted the condition that should be satisfied (it should be a number taken from the bag, which means *divisible by 4*, *"We are working with all the numbers divisible by 4"*) and the condition that was contradicted (*"4 is in the bag and 4 is not divisible by 8, 20 is in the*

---

[91] Gessenia's emergent understanding of the sufficiency of one counterexample to disprove a UAS was discussed in Section I.2.1.1, above.

[92] For details, see Section I.2.2.2 above or Appendix A10.

bag and 20 is not divisible by 8, 12 is in the bag and 12 is not divisible by 8") of St151, which was another application of her assumption aAGt2.

Figure 22 shows the development of Gessenia's assumptions about the characterization of counterexamples to a UAS and the elements or aspects of the intervention that might have supported or triggered changes[93].

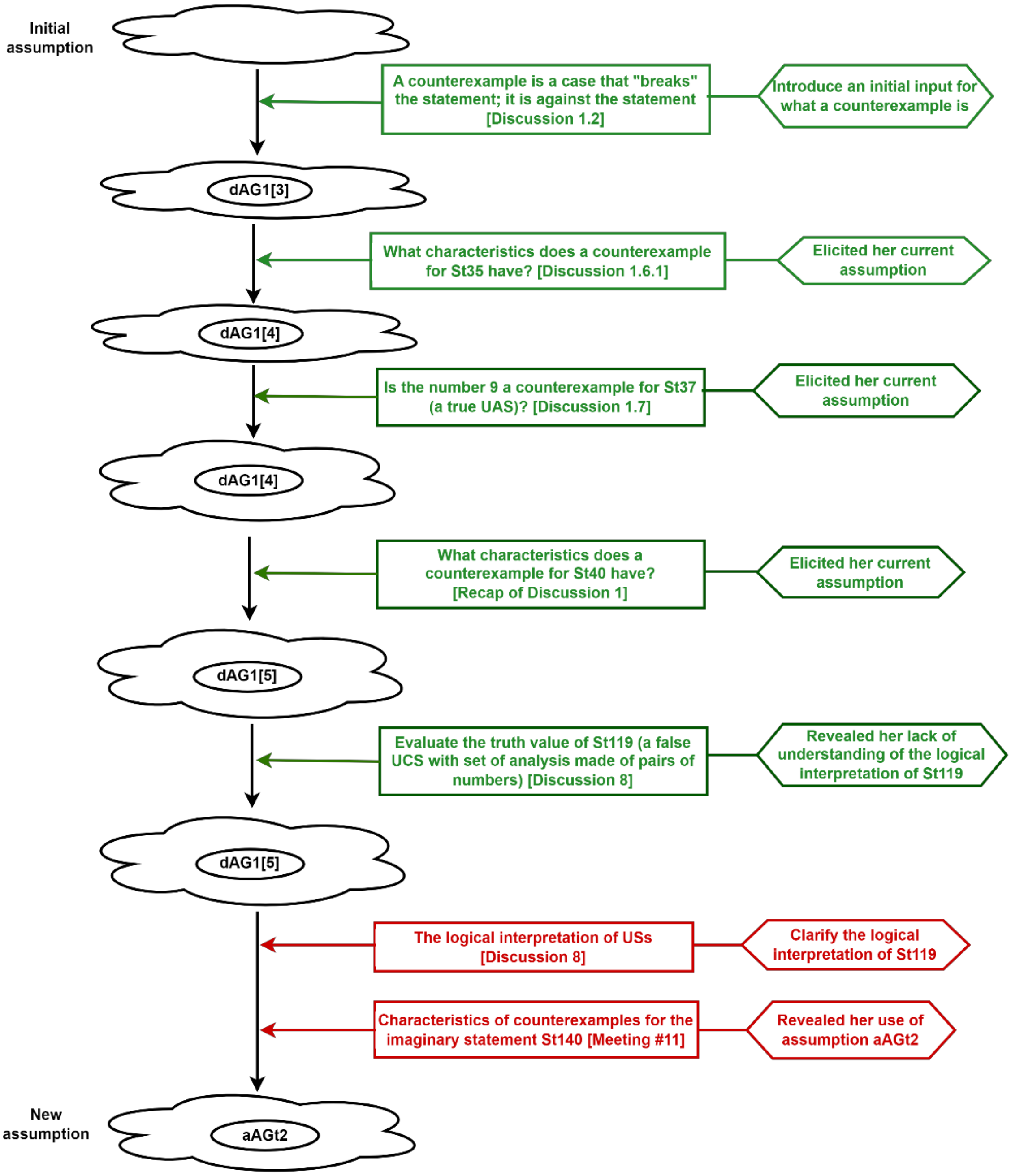


Figure 22. The development of Gessenia's assumptions about the characterization of counterexamples to a UAS (dAG1[3]: A counterexample should "break" the statement; dAG1[4]: Counterexamples should not satisfy at least one condition in the statement; dAG1[5]: Counterexamples should contradict the statement; aAGt2: A counterexample should verify the first condition, but contradict the second condition of the UAS)

93 See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

Summary of Section I.2.2

Andrea was the only teacher who began the intervention with an initial assumption about "counterexamples". Specifically, she claimed that she used the term "counterexample" in her teaching. Nonetheless, her use of "counterexample" differed from its counterpart in mathematics, which means that it was not linked to analyses of the truth value of mathematical statements. Hence, Andrea was not aware of the specific description that counterexamples should have in relation to a statement before the intervention. This was also clear for the case of Gessenia and Lizbeth, who showed they did not have any previous (conscious) understanding related to counterexamples in particular and more generally related to universal mathematical statements and how to disprove them.

In the process of determining the precise description for all possible counterexamples that would disprove a universal affirmative statement, three aspects of the intervention complemented each other to achieve this goal: the initial incomplete input for what a counterexample is, discussions with a focus on the logical interpretation of the statements and the suggested challenging examples. I suggested *irrelevant* and *confirming examples* for the UASs in discussion, which had to be discarded as counterexamples by the teachers. In the process of explaining why such challenging examples did not qualify as counterexamples, the teachers were expected to identify the conditions that counterexamples should and should not satisfy.

Lizbeth's progress towards finding the description for counterexamples began early during the intervention. Unlike Gessenia, for Lizbeth the challenging examples seen during the intervention played a vital role in that process from the start. Lizbeth could identify first the description for the possible counterexamples for specific false UASs, which suggested that she had precisely identified what in the statement a counterexample should contradict. When a counterexample for the converse of a statement was suggested as if it were a counterexample for the statement, Lizbeth found out that the order in the statement was an important aspect to consider. Notably, it played an important role when determining the characteristics of a counterexample.

On the other hand, Andrea seems to have based her reasoning on Lizbeth's "open" noticing process. Unlike Lizbeth, Andrea managed to formulate a general characterization that summarized a characterization of all possible counterexamples for a generic UAS during the intervention. When new terminology was introduced for conditional statements, Andrea only refined her characterization by replacing terms with those from the new topic. Andrea applied her new insight in two interesting contexts: first, in her search for counterexamples to a universal *negative* statement, for which a counterexample was hard to determine; second, her acceptance of a *non-minimal* justification, which led her to coin the expression "verbalized counterexample" and consider it a valid justification to disprove a UAS.

Gessenia's assumptions included her ambiguous assumption that a counterexample contradicted at least one condition in the statement, without being explicit about the exact condition that should be contradicted. Gessenia used this assumption throughout the intervention. Explicit evidence of a shift from her imprecise first assumption arose only after the intervention, during her teaching. Only by then she emphasized the specific characteristics the counterexample for the statement under analysis should have. Those characteristics were aligned with the general characterization that Andrea provided during the intervention. Later she used the same criterion to identify the description of a *hypothetical* counterexample for the case of an imaginary statement. This suggested that she used a general characterization for counterexamples.

The three teachers' teaching of false UASs and their disproving appealed to the logical interpretation of the statements and to challenging examples. Notably, they drew their students' attention to those as they attempted that their students identify the characteristics of the counterexamples for the UASs in discussion.

3. Proving of Universal Affirmative Statements

In this section I focus on the development of the teachers' assumptions related to proving universal affirmative statements. This was the main focus of Discussion 6 of the intervention for teachers; however, as I show below, previous discussions already included some debates related to this topic.

This section is split into three main sections directly linked to the framing topic of proving UASs: confirming examples and their status when proving UASs (Section 3.1); the choice and use of examples when formulating, evaluating and attempting to prove a universal statement conjecture (Section 3.2); an emergent assumption about proving UASs as the non-existence of counterexamples (Section 3.3).

These three sections outline the main issues in which the teachers were engaged in relation to proving UASs during the intervention for teachers.

3.1. Confirming examples and their status when proving UASs

In this first section I include the development of two teachers' (Andrea's and Lizbeth's) understandings.

The analysis I developed for the results in this section was based on Buchbinder and Zaslavsky's (2009) work, which includes a framework for understanding the status of examples in establishing the truth of mathematical statements. In particular, this framework encompasses the status of confirming examples when proving universal statements, which is the main focus of this section. The findings I present herein are evidence that Buchbinder and Zaslavsky's framework can be extended to cover more fine-grained aspects about the status of confirming examples when proving universal statements. The development of Andrea's understanding of the status of confirming examples in relation to the number of cases involved in a universal statement opened this possibility. On the other hand, the development of Lizbeth's understanding of the status of confirming examples when "proving" a *false* UAS that admitted confirming examples played an important role in the identification of Lizbeth's personal meanings that directly influenced her conclusions about proving universal statements.

Before the intervention the three teachers shared a common initial assumption about proving true universal statements and the status of confirming examples in this process. The teachers assumed that their use of examples counted as conclusive evidence when arguing why a universal statement that involved an infinite number of cases was true. Other researchers have shown that this is indeed a widespread assumption among students at different educational levels (see e.g., Balacheff, 1988; Chazan, 1993; Harel & Sowder, 1998; Healy & Hoyles, 2000; Zaslavsky & Shir, 2005, for details see Chapter 2, Section I.2.1). Nonetheless, the examples the teachers used with this purpose revealed differences similar to those Martin and Harel (1989) described in their research. Andrea tended to use examples that exhibited a *pattern* in order to address generality; whereas Gessenia and Lizbeth employed isolated *examples* that did not indicate any intention to prove the statement generally, similar to Balacheff's (1988) *naïve empiricism* or Harel and Sowder's (1998) *inductive proof scheme*.

We all used the words *justification* and *proof* interchangeably to refer to a mathematical proof[94]; however, since most of the tasks involved the term *justification*, the teachers mostly used that term.

3.1.1. Andrea's assumptions

The development of Andrea's understanding of the status of confirming examples to prove UASs was gradual and several changes and refinements were observed.

Table 12 gathers Andrea's assumptions related to the status of confirming examples when proving UASs. They are presented in the first column in the order they were observed. The type of assumption (whether it is an initial or intervention-based, new, assumption) is indicated in the second column.

Table 12. Andrea's assumptions about confirming examples and their status when proving UASs

Assumption	**Type of assumption (when first observed)**
bAA5: Confirming examples are sufficient to prove UASs that involve infinite cases.	Initial assumption (before the intervention)
dAA1[10]: Confirming examples are insufficient to prove.	New assumption (during the intervention)
dAA1[10]*: Confirming examples are insufficient to prove a universal statement	New assumption (during the intervention)
dAA3[1]: As long as the set of analysis is large, examples are not valid; if it is a small set, then they are.	New assumption (during the intervention)
dAA9[9]: When the statement is universal and true with infinite cases involved, an example is not a valid justification, unless it is a generic example.	New assumption (during the intervention)

Andrea's initial generalization-from-a-pattern approach (bAA5)

I identified Andrea's initial implicit assumption before the intervention and confirmed it during the first part of the intervention[95]. I observed it not only through her evaluation of others' arguments, but also in arguments of her own. Andrea implicitly assumed that by verifying a number of examples that followed a pattern, she could justify an infinite universal statement[96].

Before the intervention, during the First Exploratory Interview, Andrea displayed her reliance on the use of examples when arguing about the truth of a universal statement in two ways. First, her use of the terms "justification" and "proof" was tied to her understanding of "discovering a property". For Andrea, the "justifications" were included in a broader process that she regarded as "discovering properties in mathematics". She assumed that the later process was aligned with what is usually known as a generalization based on a pattern identification. Second, her analysis of an argument revealed the

[94] In Chapter 3, Section I.2.1 I included the meaning of (mathematical) proof that I adopted for my analyses.
[95] The first part of the intervention had a focus on the mathematical content (see Chapter 4, Section II.2.1 for details on the design of the intervention).
[96] By *infinite universal statement* I mean a universal statement that involves infinite cases; in other words, its set of analysis is infinite.

relevance she gave to confirming examples in order to guarantee a true universal statement.

During the first part of the First Exploratory Interview[97] the teachers were asked whether during their classes they discussed mathematical properties and to provide examples if that was the case. Andrea answered affirmatively and mentioned the commutative, associative and distributive properties, where only natural numbers and the multiplication and addition operations were involved. Andrea claimed that she did not teach properties, but she guided her students to "discover" them. Based on what she answered, by "discover a property" Andrea meant *verify some examples, identify and generalize a pattern based on the studied cases*. As an illustration she explained that the commutative property of multiplication was "discovered" by matching rectangular arrays, such as arrangements of soldiers in 2 rows and 5 columns and in 5 rows and 2 columns. She asserted that this was the way her students realized that the order of the factors did not change the product, and then they agreed that this was a general property for multiplication. Andrea also claimed that the justification for the property was "in the discovery process"; however, at least in the examples she provided, she did not provide or focus on the general reasons why the identified pattern held for all the numbers involved in the statement.

> Andrea: *We do not teach properties as such. We aim for the properties to be discovered. The class realizes that it holds in this situation, it also holds in this other situation, then it becomes a property. That is what I call a property. In this way they [the students] discover a property… the justification for the property is in the discovery process.*

Later, during the second part of the First Exploratory Interview[98] the teachers were requested to explain why the property *"The sum of two even numbers is an even number"* is true. First, Andrea reacted by claiming that it was the first time she had seen this property. The strategy she used afterwards to "justify" it was consistent with her *discover a property* approach. Andrea started by explaining that numbers are distributed in the number line as odd and even (beginning with 1), one after the other. She asserted that when two even numbers were added, like 2 and 6, the result was also an even number; however, she did not explain why that would always be the case. If the numbers were bigger (e.g., 10 and 4), her explanation relied on the unit digits to conclude that the sum would be even (*"14 ends in 4 that is even"*). To reach that conclusion, Andrea used her implicit assumption that even numbers have unit digits that are even. Andrea's explanation exhibited a lack of focus on discussing the general reasons for why a pattern held beyond the examples she explicitly explored. The examples she tested did not exhibit a pattern from where to extract the missing general reasons either.

Another task involved the evaluation of a given argument. The task was framed as a classroom episode in which a student, Rodrigo, came up with the conjecture *"If we multiply any natural number by 5, we will get a number with the units-digit equal to 0 or 5*". Another student in the same class, Sebastian, made the claim that Rodrigo's conjecture was true *"because if you multiply 11 by 5, you get 55, which ends in 5. If you multiply 24 by 5, the result is 120, which ends in 0. So, yes, you will always get what you just said"*. The teachers were asked to answer two questions that followed the classroom episode:

---

[97] The first part of the First Exploratory Interview consisted of a semi-structured interview. The questions I used for that interview can be seen in Appendix EI1-P1.

[98] The second part of the First Exploratory Interview consisted of a task-based interview. The tasks I included in that interview can be seen in Appendix EI1-P2.

a) What score would you give to Sebastian's justification for Rodrigo's "mathematical discovery"? Think of scores from 1 to 5, where 5 is the best score. Explain why you gave such score and not a different one;

b) Do you think that Sebastian's justification guarantees that Rodrigo's "mathematical discovery" is true? Why? How do you know?

Andrea gave Sebastian's justification the highest score available (5) because, as she explained, Sebastian verified Rodrigo's mathematical discovery with examples and he did not only accept it passively.

Andrea: *I would give him a 5, because he is verifying it… when he provides examples. Besides, he is putting this into words. He has thought of an example for what Rodrigo said, and he is not only accepting this, but he is indeed verifying it… He [Sebastian] is verifying that indeed what he heard from his fellow Rodrigo does follow.*

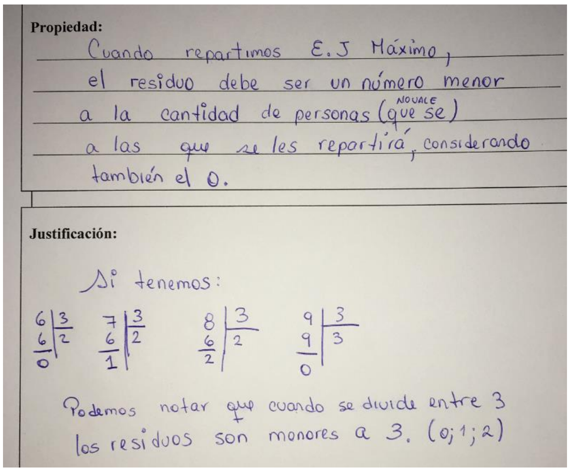
Andrea also answered affirmatively that Sebastian's justification guaranteed that the conjecture was true. Nevertheless, she clarified that even though Sebastian verified Rodrigo's discovery, this was not enough to show why Rodrigo's discovery was true. She considered that Sebastian needed to provide more examples in order to do so.

Andrea: *Yes, he [Sebastian] is verifying this. He does not necessarily show why, because he would need to provide more examples… three or four more so that he can certainly say that yes, this is true, it does verify what Rodrigo claims.*

Based on Andrea's evaluation of the argument in the classroom episode two things could be noted: (1) Andrea valued the fact that the student did not only accept a classmate's claim, but that he was eager to *verify* (in the sense of *confirm*) it; (2) Andrea's request for the use of more examples to show why the mathematical discovery was true indicated her own at-that-moment expectation of what needed to be shown in order to guarantee that a universal statement was true. Andrea expected to see a number of confirming examples she would have admitted as a "justification".

The use of Andrea's initial assumption bAA5 was also evident during the intervention for teachers. Specifically, during the first part of the intervention, focused on the content of division and divisibility, the teachers were asked to solve an activity individually in order to engage in a further whole-group discussion afterwards. Activity 5 included two tasks that were focused on discarding impossible values for the remainder, either when it was presented as the number of objects left in a FWM-distribution (in Task 1; e.g., it is not possible that we have 5 objects left in a fair, whole and maximal [FWM] distribution of objects among 3 people)), or explicitly as the remainder in a division (in Task 2; e.g., it is not possible that a division by 4 leaves a remainder of 6)[99]. This activity was continued with another individual activity that requested the teachers to identify a property[100] based on their solutions to Activity 5 and our discussion. They were also asked to provide a justification for it. This was the first opportunity the teachers had to formulate and justify a property during the intervention; however, no criteria for writing down justifications were explicitly discussed before. Figure 23 shows Andrea's first property and her respective justification.

[99] For details on the activity and the tasks that were part of it, see Appendix A5.
[100] We used the word "property" instead of "theorem" mainly because the teachers were more familiar with it.

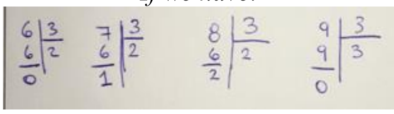

Propiedad:
Cuando repartimos E.J Máximo, el residuo debe ser un número menor a la cantidad de personas (NO VALE que se) a las que se les repartirá, considerando también el 0.

Justificación:
Si tenemos:

Podemos notar que cuando se divide entre 3 los residuos son menores a 3. (0;1;2)

Figure 23. Andrea's formulation of property 1 and her justification for it, based on Activity 5 (Original version in Spanish; an English translation to the right)

Property 1:
When we distribute (objects) in a FWM-way, the remainder must be a number smaller than the number of people to whom the objects will be distributed, considering number zero too.
Justification:
If we have:



We can notice that when we divide by 3 the remainders are smaller than 3 (0; 1; 2)

Andrea's formulation of her property was general. Her justification included four divisions by 3. She might have seen the number 3 as a generic number of people (or divisor) that she used to explain its relationship with the possible values the remainder could take. However, her justification did not include any mathematical reasons why the remainder could not be equal to or more than 3. Unlike other approaches of teaching division, the FWM-distributions approach does not include the relation between the remainder and divisor (the remainder should be smaller than the divisor) as a pre-existing condition that should rule the division process. Instead, the FWM-distribution approach allows reasoning to explain why it makes sense to include this as a property of divisions.

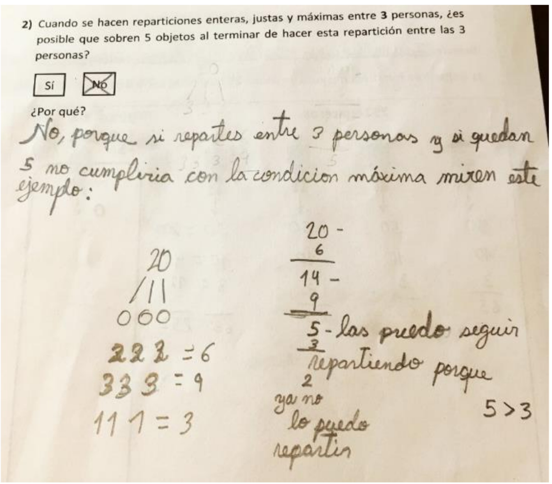
Andrea's justification relies on the use of a pattern that she seems to have identified with the use of four examples (6, 7, 8 and 9 divided by 3, respectively), but apparently, she did not see the need to explain why 3 or a greater number are not valid values for the remainder when dividing by 3. Her justification revealed again her view of a justification being included in the discovery process of a property, which is a generalization based on a pattern.

Andrea's initial use of examples in her "generalization-from-a-pattern approach" to prove a UAS was structured. Moreover, the fact that Andrea relied on patterns and generalizations from those patterns suggests that the evidence that she used to prove UASs was empirical, as it did not encompass all the cases involved in the statement. Likewise, her evidence did not include the general reasons why the pattern would follow beyond the cases she analyzed. Hence, her approach did not lead to proof in a mathematical sense.

Andrea's extreme assumption dAA1[10]

Reorienting attention to general reasons through the analysis of others' arguments

An emergent deviation from Andrea's use of "generalization-from-a-pattern" approach to justify a general claim surfaced, still during the first part of the intervention. Before Activity 6 the teachers' attention was drawn to the analysis of arguments. All the arguments concerned the relation between the remainder and the divisor. The arguments were presented as classroom episodes (e.g., see Appendix CE8.2) and actual students' arguments for analysis. An illustration of one student's argument that the teachers had to evaluate is shown in Figure 24.



2) When fair, whole and maximum distributions are developed among 3 people, is it possible that 5 objects are left once the distribution is finished?
Yes ~~Not~~

Why?
No, because if you distribute among 3 people and there are 5 (objects) left, it would not satisfy the maximal condition. Look at this example:

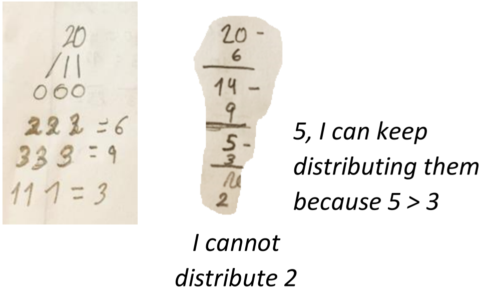


Figure 24. A student's argument for why 5 objects cannot be left in a FWM-distribution among 3 people used for discussion during the first part of the intervention for teachers (the original version in Spanish on the left; an English translation on the right)

This shift away from using examples was evident in Andrea's solution to Activity 6 during the first part of the intervention. In that activity the teachers were asked to find the minimal and maximal remainder as the divisor changed in a division (the examples of divisors included were: 2, 3, 10 and 17). They were also requested to complete sentences like "When it is a division by 58, the maximal reminder is _____" so that the sentence is true and asked to explain why the remainder cannot be a bigger number[101].

Andrea's argument did not depend on the use of an example; it focused on a more general issue, in contrast to her first justification in Figure 23. Notably, she focused on why the remainder could not be bigger than or equal to the number of people (the divisor). Andrea's justification changed in the sense that it was no longer based on examples that exhibited a pattern without really explaining why the pattern held beyond the examples provided. Andrea's immediate reaction to Lizbeth's suggestion to use an example to prove the statement was to claim that her justification would not be complete yet. Andrea's experience in reading others' arguments seems to have directly influenced the way she reasoned the relation between the remainder and the divisor in a division in a general way. For example, in Figure 24 the student explains why in a FWM-distribution no more objects than the number of people could be left. This idea was used by Andrea, who seemed to have been persuaded by the student's answer as she used its main idea to justify her own general property.

Analyzing whether certain arguments proved a UAS seems to have significantly contributed to a first shift in the form of reasoning Andrea used during the first part of the intervention. The teachers analyzed third graders' arguments, which mainly involved the use of text-form answers, as well as dialogues between a teacher and her class (classroom episodes). For example, when we discussed the argument in Figure 13 we agreed that the student provided the example of 20 to illustrate a more general idea that she had expressed as a verbal explanation before and not because the example was the most important part of the argument. Presumably, from this sort of experience Andrea assumed that a general argument is usually presented as narratives or have a text-form (most properties and

[101] For details on the Activity 6, see Appendix A6.

conditions we used were presented in that form of expression). Besides, she seems to have learned as well that arguing why a pattern was true based on some confirming examples without explaining the reasons for why such pattern will continue to hold was not enough. This means that at that point, some norms for justifications were implicitly integrated in our discussions.

Andrea's emerging assumption that confirming examples alone were not sufficient when proving a universal statement might have been reinforced by a clarification she requested. At a late stage of the first part of the intervention Andrea asked about the difference between a *justification* and a *verification*, which suggested that she had already begun to notice differences between them. Most likely, her question was a response to the examples of what we accepted as justifications for true universal statements (properties) the teachers deduced, which were basically inferences from definitions (i.e., direct proofs). Given that this was not the main aim of the first section of the intervention, the teachers were provided some incomplete ideas about this distinction. In short, I told them that a *verification* is a confirmation that the statement holds with the use of examples, whereas a *mathematical justification* goes beyond a verification as it might involve complex ideas. I also said that details would be discussed later during the second part of the intervention.

As a result of Andrea's experiences during the first part of the intervention, Andrea discarded the sufficiency of confirming examples to prove a universal statement during the second part of the intervention for teachers as I illustrate next.

During the second part of the intervention Andrea explicitly referred to the status she attributed at that moment to confirming examples when proving UASs: Confirming examples are insufficient to prove (her assumption dAA1[10]). For example, during the Recap of Discussion 1 the teachers were asked to provide examples of false universal statements. Andrea provided the statement *"All numbers divisible by 3 are even numbers"*. Lizbeth reacted by claiming that the number 6 proved the statement (see the development of Lizbeth's assumptions in Section I.3.1.2 below). Andrea rejected Lizbeth's assertion and corrected Lizbeth by suggesting that if it were a true universal statement, it would not be enough to verify an example to prove it. Andrea's response to Lizbeth's claim revealed the use of her assumption dAA1[10].

Andrea showed again her use of her assumption dAA1[10] during Discussion 3.2. The discussion was put forward in the context of a student named Pepito and the conjecture that he formulated: *"All palindrome numbers are divisible by 11"* (St72). Pepito claimed that his conjecture was true because of four "big" palindrome numbers he verified with his calculator. The teachers were engaged in a debate about whether Pepito's conjecture was true and whether his justification was valid. In this discussion Gessenia showed her approval of Pepito's verification of the four confirming examples as a way to prove St72. Andrea reacted to Gessenia's response by pointing out that when a statement was true, examples were not sufficient to prove it, suggesting that the four palindrome numbers Pepito verified did not prove his conjecture.

Episode 12 (below) shows a dialogue developed during the continuation of Discussion 3.2, where Andrea expressed the status she had established for "examples".

From Episode 12 I extract three aspects about Andrea's assumptions related to examples and their role and status when proving UASs. First, Andrea's use of the term "examples" sometimes referred to *confirming* examples, and other times to *contradicting* examples. During the first part of Discussion 3.2 Andrea used the word "examples" to mean *confirming examples*. She suggested that the "examples" that Pepito provided did not prove his conjecture and all those examples were supporting examples. On the other hand, in turn 10 of Episode 12 she used the term "examples" to mean *counterexamples* (*"So, the examples will help them, BUT as a justification, ONLY for false universal statements"*). At this point of the intervention Andrea was already aware that counterexamples were sufficient to disprove a universal statement (see Section I.2.1.2 above). In consequence, she used the term "examples" in a broader sense to refer to (either verifying, or contradicting, accordingly) "computations" in accordance with the statement in discussion.

Episode 12

Turn	Who	What
1	I	*However, what happen with what Pepito says? Is his justification valid? How do you understand what a valid justification is?*
2	Gessenia	*I understand it like, he-, I mean, his examples are correct.*
3	Lizbeth	*He is claiming, he is confirming that what he has done is true.*
4	I	*When I ask if a justification is valid, I am asking whether the justification indeed PROVES that the statement is true, since that is what Pepito claims in this case.*
5	Andrea	*That means, if it [the justification] is sufficient.*
6	I	*That means if it is sufficient, exactly. Is his justification sufficient? Does it guarantee what he is claiming? He claims that his conjecture is true. Does what he shows guarantee that his conjecture is true?*
7	Lizbeth	*No, because he did not analyze all palindrome numbers.*
8	Andrea	*I believe that examples are valid when the given statement is false and universal.*
9	I	*Let's see, I do not discard… I mean, it is useful that children begin by exploring examples, right? Because obviously they are not going to jump directly to something abstract. Most likely, they are going to begin by analyzing examples and they are going to say, let's see, I'm going to test 121, with another one, and in the case of Pepito, maybe he just had bad luck and chose the examples that are indeed divisible by 11. But there might be one student who finds the case that, not, but 131? 101? Those do not hold. Aha! Then the conjecture is not true.*
10	Andrea	*Besides, they [the students] don't know if it [the conjecture] is true or false. So, the examples will help them, BUT as a justification, ONLY for false universal statements.*

Second, she distinguished between the role that examples might play and their status when proving universal statements. On one hand, Andrea recognized the role of examples as supporting the exploration and evaluation of a conjecture in order to determine whether it is true or not (*"Besides, they don't know if it [the conjecture] is true or false. So, the examples will help them"*, turn 10). On the other hand, she delimited the role of examples and their status (*"BUT as a justification… ONLY for false universal statements"*, turn 10). In other words, Andrea expressed that examples were useful for exploring the statement in either case, whether the statement was true or false; however, in order to justify the statement, examples were *only* valid when the statement was false and universal; that is, to disprove a US.

Third, Andrea assumed that examples were only valid justifications when refuting universal statements. Andrea asserted that examples constituted valid justifications, in the sense of "sufficient justification" (see turn 5), only when disproving universal statements (see turn 10). As a result, her claim discarded cases where examples could prove true universal statements (her assumption dAA1[10]).

Andrea's first refinement of her assumption dAA1[10] (dAA3[1])

After Discussion 3.2 the teachers were asked to share their new insights related to conjectures, justifications and mathematical truths. Episode 13 is the part of the dialogue that has a focus on Andrea's insights. In this episode Andrea shows evidence of a reconsideration of her assumption dAA1[10] that examples cannot prove a true universal statement in order to conclude that there may be cases of true universal statements for which the verification of examples may be indeed sufficient justification.

Episode 13

Turn	Who	What
1	Andrea	*I believe that for a conjecture to be true, I mean, if I believe that my conjecture is true, I do not justify it with examples. I have to justify it in a different way, but not with examples, because only examples do not-*
2	I	*In order to be true?*
3	Andrea	*Right, I mean, to justify it, when a conjecture is true, that is if I am sure that my conjecture is true, I do not justify it with examples. Because I'm realizing that examples only work for false conjectures.*
4	I	*Let's see, I already have an example-, but, when you say conjecture, what do you mean? Do you mean any conjecture?*
5	Andrea	*Statement.*
6	I	*Any statement?*
7	Andrea	*Sure.*
8	I	*Ok then, I will give you one example.*
9	Andrea	*Ah, no, universal!*
10	I	*Could you [to Gessenia and Lizbeth] think of a true universal statement that can be justified with examples? Because Andrea claims that it is not going to be possible to find any. I say, that is not exactly true… I have a counterexample for your [Andrea's] claim.*
11	Andrea	*Of course, as long as the group is large; of course, because if it is a small group, sure, yes. For example, if I say, um, the numbers of one digit, right? … Yes, it is possible. I mean, it ["examples are not a valid justification"] is like a recommendation if I say that it is a big number [of elements involved], that is, a big set, if it is a large set.*
12	I	*Ok, Ok. (laughter) Aha! But at the beginning you said that it was not possible, and it is actually possible.*
13	Andrea	*Yes, it is possible. I mean, it ["examples are not a valid justification"] is like a recommendation if I say it is a big number [of cases involved], that is, a big set, if it is a large set.*
	[Attention is drawn to the statement *"All one-digit numbers[102] divisible by 6 are divisible by 3"* and on determining: whether it was a universal statement, the set of analysis, whether it was true and its justification]	
14	I	*Then, was it sufficient that with these cases (0 and 6) I prove-, look, it is universal and true. Were those verifications sufficient to prove that the statement is true?*
15	Lizbeth	*Yes, because the only one-digit numbers divisible by 6 are 0 and 6, and 0 and 6 are also divisible by 3.*
16	Andrea	*Yes, because those are the only ones.*

[102] Our "framing set" was the set of natural numbers. Discussion 1.7 included the analysis of a similar statement (*"The numbers divisible by 6 are divisible by 3"*).

Andrea began her participation by relying on her assumption dAA1[10] (see turns 1 and 3). She was then challenged to make her claim more precise, in regard to different aspects: whether she rejected examples to justify true or false conjectures (turn 2); what she meant by conjecture (turn 4); whether by conjecture she meant a universal or existential statement (turns 6 and 8). To her refinement of her assumption dAA1[10] where she specifies that examples cannot prove a universal statement I call Andrea's assumption dAA1[10]*.

In turn 10 I asserted the existence of universal statements that can be verified using examples (*"... Andrea claims that it is not going to be possible to find any. I say, that is not exactly true... I have a counterexample for your [Andrea's] claim."*). It was clear that Andrea had not anticipated any such case on her own, but she quickly jumped into the discussion in order to modify her initial assertion by focusing on the "size" of the set of analysis. She refined her assumption to specify that when the statement involved a "small group" in the sense of a small number of cases, then verifying examples could be used as a determining approach to prove such a statement (turn 11). Andrea also stated that, in the case of a true universal statement that involved a "big" number of cases, examples were not sufficient justification.

I called her new assumption dAA3[1]: *As long as the set of analysis is large, examples are not valid; if it is a small set, then they are.* In this way, Andrea emphasized the size of the set of analysis as an important factor to determine whether confirming examples proved or not a true universal statement. Thus, the size of the set of elements involved (the set of analysis) in the statement mattered and she referred to the set of numbers of one digit (turn 11) to illustrate her insight.

Andrea's safe true conjecture: Her use of assumption dAA3[1]

After the Recap for Discussions 3.1 and 3.2 the teachers were asked to solve a task that consisted in identifying a pattern, formulating a conjecture and evaluating the truth value of the conjecture[103]. The teachers came up with the conjecture:

"For every natural number n *(different from zero), when substituted in the formula* $1 + 1141n^2$*, the result is not a perfect square number"*

Lizbeth had already verified the first ten non-zero natural numbers plus the numbers 22, 66, 120 and 1555. For all those values the conjecture held. I asked the teachers if they were sure that the conjecture was a mathematical truth. Andrea acknowledged that she was uncertain about whether the conjecture was true for all non-zero natural numbers, where the conjecture was defined. Because of her uncertainty, she suggested a variation of the conjecture formulation with the aim of achieving truth.

> Andrea: *[I could be sure that the conjecture is a mathematical truth if] I created another conjecture... But in order for her conjecture to be true, then let's only take and refer to the natural numbers from 1 to 10. Let's substitute them there, right?*

Andrea restricted the set of analysis (the set of non-zero natural numbers) of the conjecture to a finite set of numbers (the set of the first ten non-zero natural numbers) in order to guarantee that the conjecture was a mathematical truth. She was sure that the conjecture was true for her suggested new set of analysis as they had verified that the

[103] For details on the task content, see Appendix D3, or details on its context, see Section I.3.2 below.

conjecture was true for those cases. With her recommendation Andrea revealed her understanding that a universal-statement conjecture defined in a finite set of elements is guaranteed to be true through a verification of all the cases involved in it, in contrast to those that are defined in an infinite set of cases. Thus, Andrea's strategy calls for a set of analysis where she is already aware and certain that the conjecture is safe to be a mathematical truth.

A similar episode is traced back to Discussion 6.2.3 where the teachers were shown two (true) universal statements in parallel, one that involved infinite cases (*"0 is divisible by any natural number different from zero"*) and the other involved a finite number of cases (*"All natural numbers smaller than 5 are smaller than 7"*). One of the tasks asked the teachers for the type of justification that was expected to guarantee the truth value of a true universal statement[104]. Andrea's answer revealed again the use of her refined assumption dAA3[1].

> Andrea: *It depends ... if it is finite, the set of analysis is finite, it could be justified with examples. But if it is INFINITE, I would need to do it more theoretical or justify it generally.*

Andrea's accurate answer again exhibits her attention to the size of the set of analysis and how this affects the kind of argument to be given in order to prove a universal statement, and hence, her use of assumption dAA3[1].

Andrea's second refinement of her assumption dAA1[10] (dAA9[9])

The case of generic examples turned to be more challenging for Andrea, who initially did not consider them to be valid justifications to prove UASs with infinite cases involved.

Towards the end of the intervention the teachers engaged in a discussion (Discussion 9) with a focus on whether or not a *generic example* could prove a UAS with infinite cases involved. Andrea was very skeptical about this. At first, her main concern was the use of generic examples, since Andrea regarded them as merely confirming examples; that is, she could not "see the general in the particular" (Mason & Pimm, 1984). She expressly reminded Lizbeth and Gessenia about Discussion 3, where a false universal statement was disproved by an "extreme" counterexample[105] (see above, and Section I.3.2 below for details). For Andrea, there was no guarantee that a statement could be proved to be true through showing examples that verified the statement and explaining something in a general way through the example; otherwise, we would not have had a case like the statement in Discussion 3, for which a counterexample was not expected at all and it was actually not accessible to us by only using paper and pencil, or even calculators.

Andrea's response exhibited her initial stance about generic examples, that they only verified the universal statement, the same way as confirming examples would do; however, those confirming examples did not count as conclusive evidence, that is required to prove a universal statement. Her doubt here is similar to one that Chazan (1993) suggested needed more attention. Chazan asked whether it was possible that, as students learn to doubt geometric measurements when using tools like a ruler or a protractor, they might also begin to doubt deductive proofs. Here we see that Andrea was

---

[104] For the task content, see Appendix D6.

[105] "Extreme" counterexample in the sense that it was a very large number that was found to be a counterexample for the statement *"For every natural number n (different from zero), the expression:* $1 + 1141n^2$ *never produces a perfect square"* with the use of computers.

skeptical about the sufficiency of examples to prove and as a result she was reluctant to accept that generic examples could prove universal statements with infinite cases involved. This may be a wider issue related to examples and proving. As students become skeptical about empirical arguments as valid methods to prove, might they become skeptical about generic arguments that qualify as proofs?

The subsequent discussion around the acceptance or not of generic examples to prove universal statements was based on some observations I have discussed elsewhere (Reid & Vallejo-Vargas, 2018): the need to include evidence of awareness of generality (explicit claim that the pattern follows for all cases involved in the statement) and mathematical evidence of reasoning (the mathematical evidence that explains/shows why the pattern follows). Considering the case suggested by Andrea from Discussion 3, three aspects were emphasized: (1) the importance of an explanation accompanying the generic example(s) that points out the general mathematical reasons behind the generic example(s); (2) a mathematical abstract proof could be constructed by basically following the steps involved in the generic argument; (3) the case she made reference to from Discussion 3 was actually false and as a consequence it was impossible to find general reasons through the use of confirming examples to prove the conjecture.

These three aspects seem to have permeated the development of Andrea's assumptions. During Meeting #10, after the intervention, she acknowledged the case of generic examples as a possible way to prove universal statements (her assumption dAA9[9]) for the case of an imaginary statement. During this meeting Andrea explicitly accepted generic examples as a possible valid justification for true universal statements that involved infinite cases. The teachers were given the imaginary universal statement *"All numbers bigger than 5 are RAINBOW numbers"* (St139). They were requested to answer what kind of mathematical evidence was sufficient to show in order to guarantee that St139 is true. Andrea considered that an example was not sufficient. She added that a property could be used to provide support or prove the statement. An example was not enough in her view, because *"the set of analysis is infinite"*, unless – as she pointed out – it was a *"general example"*, by which she meant a generic example.

The development of Andrea's assumptions about proving universal statements and the status of confirming examples when proving those statements can be seen as a sequence of refinements that included distinctions in terms of the number of cases involved in the statement and the use of generic examples. The case of Andrea shows the precisions that need to be paid attention to so that overgeneralizations are not made. It is crucial that teachers become aware of special cases of universal statements where confirming examples may be indeed conclusive evidence to prove them.

Andrea's Teaching

Andrea's teaching showed her attention to three elements that allowed her to put forward whether an example was sufficient evidence to guarantee that a UAS was true: first, the universality of the statement in discussion; second, the set of analysis and aiming at the elements in it; and third, the number of cases involved in the statement. This shows her consistency with the insights she gained during the intervention for teachers.

For example, during her teaching of Session 13 her students were asked to solve Activity 10[106]. The first task in the activity was contextualized as a certain number of marbles that

[106] For the content of the activity, see Appendix A10.

were hidden in a box. Even though the number was unknown, it was given that the number of marbles could be evenly distributed among 6 people without having marbles left over after the distribution. The question the students needed to answer was whether the same number of marbles, when evenly distributed among 3 people, would also leave zero marbles once the distribution was over and to explain why. This means that the students were indirectly expected to identify that *"If a number is divisible by 6, then it is divisible by 3"* and prove it. They were requested to solve the activity individually first and then engage in a whole-class discussion. During the working-alone time, Andrea prompted the students to notice three things about the problem: the number of marbles inside the box was not explicitly given or known; even though the number of marbles was unknown, it could be exactly distributed among 6 people; the question was whether the number of marbles could be exactly distributed among 3 people as well. During the whole-class discussion the class tested some possible numbers of marbles that could be inside of the box. Andrea put forward an irrelevant example (14 marbles) and whether it could be a number of marbles inside the box, which is evidence of her attention to the set of analysis and the elements in it, and her desire to draw her students' attention it. Nicholas, one of the students conjectured that *"If a number is divisible by 6, then it is divisible by 3"* and explained why. Andrea saw that Nicholas' argument was general, but she noticed that many students did not follow Nicholas' reasoning, so she asked Nicholas further questions (e.g., *"and why is it important that 6 is two times 3?"*). She then reminded her students about the particular cases they had tested (0, 6 and 12). Her goal consisted in lowering the difficulty of Nicholas' general argument to a more comprehensible level by using the examples on the whiteboard as generic ones that allowed the students to see the general reasons from the particular. She first drew attention to the insufficiency of verifying some of the examples involved in the conjecture in order to guarantee its truth. She asked whether the supporting examples they had would show that the conjecture was true and emphasized the general nature of the conjecture.

> Andrea: *So, is it true or not that if a number is divisible by 6 then it is divisible by 3? And why? ... Look, because in here [the whiteboard] we only gave three examples, but kids, how many are the numbers that are divisible by 6? ... if I want to validate this [the conjecture], I would need to verify ALL those examples.*
>
> *Student: Wow! that would take all my life!*
>
> *Andrea: Exactly, but I am not going to do that, then I have to EX-PLAIN.*

The student's comment suggested an impossible approach that allowed Andrea to point out that an explanation was expected instead. With her subsequent feedback Andrea complemented what kind of explanation she had in mind.

> Andrea: *But kids, only this (she pointed to 12 on the whiteboard), only this is not sufficient. This is an example from all the infinite [examples]. You have to tell me something more general.*

In addition to her expectation for a general explanation, Andrea also emphasized that a confirming example was not sufficient in order to prove that the conjecture was true (her assumption dAA1[10]). In order to support the students' understanding, she emphasized the *infinite* number of examples they would have needed to verify.

This shows Andrea's focus on the generality of the statement as well as on the number of cases involved in it. Furthermore, she drew attention to the inefficient strategy that would entail to verify all the infinite cases involved. She exhibited her expectation for a general explanation instead.

During her teaching of Session 14, Andrea used her refined assumption dAA3[1] that as long as the set of analysis is large, examples are not valid; if it is a small set, then they are. In particular, her students had a sequence of tasks that involved quantified statements and their truth value. Specifically, the third task entailed a universal statement with a small finite number of cases involved.

Task 3: *You are given the numbers 0, 3 and 9. Now answer: Is it true that all these numbers are divisible by 1?*

Notice that this is a case of a true implicit universal statement and its minimal justification consist of a verification for each of the three cases involved in it.

> Andrea: *When I am told "all these numbers", what does it refer to? ... Here [task 3], which one is my set of analysis? ... Do we also have infinite possibilities here [in contrast to task 1]? ... Here my set of analysis only has three elements... There are only three possibilities that I am going to analyze, that's right... I can verify them all.*

As in her teaching of Session 13, Andrea's feedback focused on the following aspects: (1) the universal expression "all these numbers" and its interpretation according to the statement; (2) the set of analysis; (3) the number of elements in the set of analysis; (4) a contrast between the number of objects involved in a previous task (which was infinite) and task 3 (which was finite); (5) the verification of the three cases involved in task 3 as a sufficient justification.

Having only three elements in the set of analysis allowed Andrea engage her class in a discussion about the size of the set of analysis and the kind of justification that was valid as a consequence. Her feedback was evidence of her current understanding about proving universal statements, which in this case exhibited the use of her assumption dAA3[1].

It is interesting that Andrea's feedback was followed by a student's spontaneous question.

> Student: *Miss, if there were 100 numbers, would I have to verify them all too?*

Andrea answered negatively and suggested him that, even though that was possible, it was better to look for another approach instead, since 100 was still a big number to deal with. The student's question revealed that Andrea's feedback encouraged the student's further thinking of a case that might be seen as in-between the analyzed cases, since it did not refer to an infinite or a small finite number of cases. The episode ended with the class verifying whether each of the three numbers (0, 3 and 9) were divisible by 1 in front of the class.

Figure 25 shows the development of Andrea's assumptions about confirming examples and their status when proving UASs and the elements or aspects of the intervention that might have supported or triggered changes[107].

[107] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

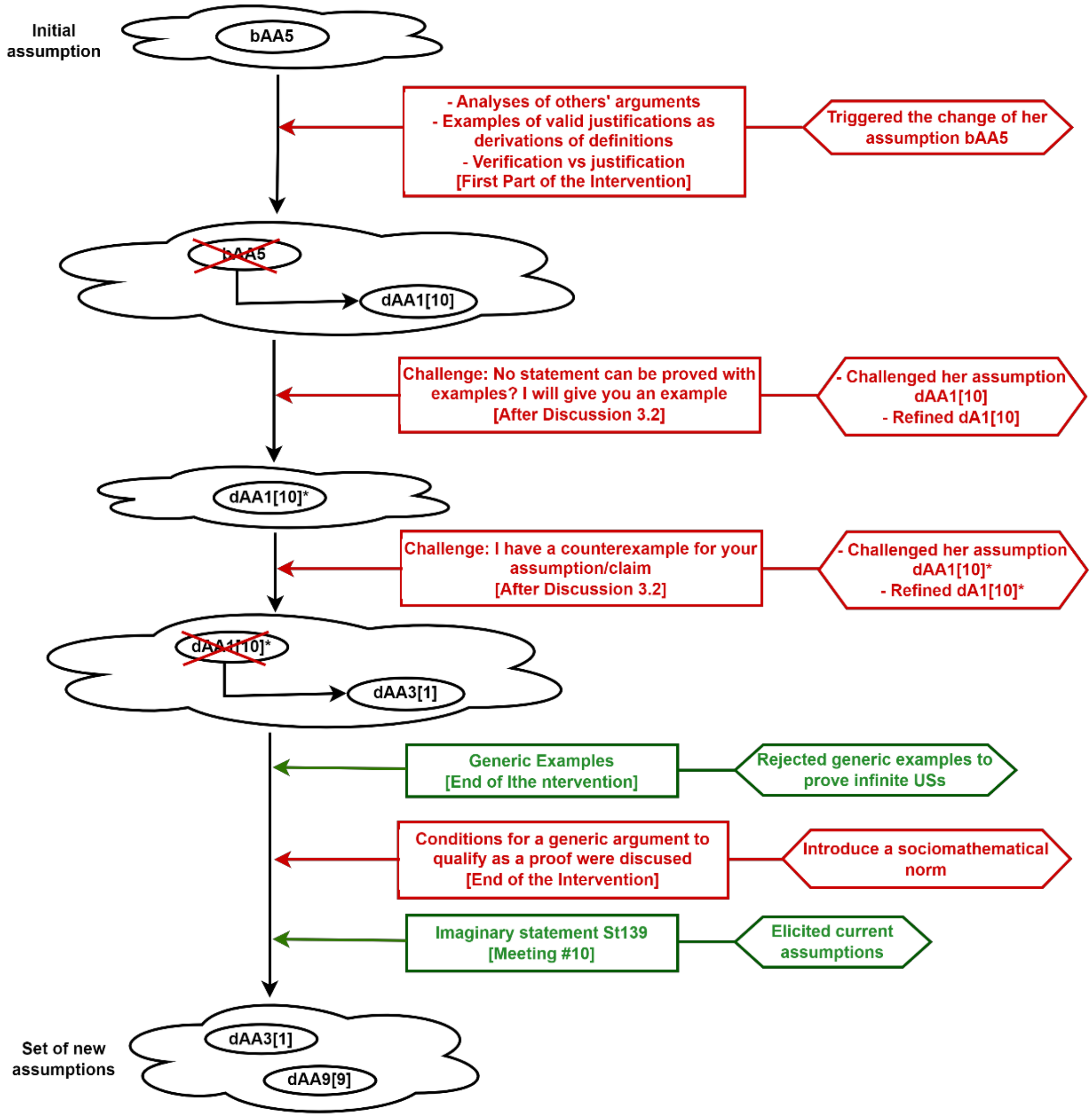


Figure 25. The development of Andrea's assumptions about confirming examples and their status when proving UASs (bAA5: Confirming examples are sufficient to prove UASs that involve infinite cases; dAA1[10]: Confirming examples are insufficient to prove; dAA1[10]: Confirming examples are insufficient to prove a universal statement; dAA3[1]: As long as the set of analysis is large, examples are not valid; if it is a small set, then they are; dAA9[9]: When the statement is universal and true with infinite cases involved, an example is not a valid justification, unless it is a generic example)*

3.1.2. Lizbeth's assumptions

Lizbeth began our discussions with an empirical scheme for proving UASs, which led her to assuming that confirming examples sufficed to prove UASs. She assumed that confirming examples justified that universal statements were true (her initial assumption bAL1).

Terms like *guarantee*, *sufficient to guarantee*, *justify* and *prove* were introduced and used during the intervention as synonyms. Our discussions had a focus on supporting the teachers' awareness of the logical interpretation of UASs[108]. Notably, this played an important role on Lizbeth's interpretation of these terms when proving (and disproving) was involved. As I show below, the development of Lizbeth's assumptions revealed a distinction she made between the terms *justify/prove* and *guarantee/sufficient to guarantee*. This was observed when she analyzed a false UAS that admitted confirming examples, as she separated her observations by claiming that: *confirming examples "justified" that the UAS was true, even though those examples did not guarantee that the UAS was true*. That is, she interpreted *"justify"* as *"confirm, verify and support"* and *"guarantee"* as *"necessarily show"*.

Lizbeth expressed her emergent understandings through her way of using these key terms and her preferences for some words over others.

Table 13 includes Lizbeth's assumptions related to confirming examples and their status when proving UASs.

Table 13. Lizbeth's assumptions about confirming examples and their status when proving UASs

Assumption	Type of assumption (when first observed)
bAL1: Confirming examples can *justify* that a UAS that involves infinite cases is true.	Initial assumption (before the intervention)
dAL1[3]: Confirming examples *prove/justify* a false UAS that admits confirming examples with infinite cases involved is true, though they do not necessarily *guarantee* that the UAS is true.	New assumption (during the intervention)
dAL3[1]: Verifying some examples does not *guarantee* that a UAS is true. All examples need to be verified in order to do that.	New assumption (during the intervention)
dAL3[2]: In order to *justify* that a statement is true, we should verify each case involved in the statement.	New assumption (during the intervention)

Next, I illustrate Lizbeth's assumptions, which have a focus on her use of "prove/justify" and the way it has an impact in the development of her assumptions related to proving UASs.

True UASs with infinite cases involved

Lizbeth's initial approach to showing the truth of UASs exhibited her *naïve empiricism* (Balacheff, 1988). Her way of proving UASs did not address the issue of generality and she did not provide general reasons to conclude that a UAS was true. This was observed not only in her reliance on the verification of confirming examples to prove (true) UASs, but also in her validation of others' arguments. The latter also revealed her tendency to consider as valid those arguments with the same characteristics as hers, where no intention to examine generality was evident (her initial assumption bAL1).

[108] For details on what I mean by the logical interpretation of single-quantified statements, see Chapter 3, Section II.1.

An illustration of the way Lizbeth used her naïve empiricism was clear during the First Exploratory Interview, before the intervention. The task requested her to prove that *"the sum of two even numbers is an even number"*. Lizbeth relied on the verification of a few specific confirming cases to conclude that the statement was, as she put it, *always* true. She did not only use a small number of concrete supporting examples to show that she understood the statement, but she also acknowledged the truth of the infinite universal statement because of those confirming examples. A similar criterion she used when evaluating given arguments before the intervention and during the first part of the intervention. She accepted arguments that only contained one or two confirming examples as valid proofs for universal statements.

False UASs that allow confirming examples: Establishing meaning

Discussion 1 of the second part of the intervention contained evidence of Lizbeth's emergent establishment of meanings for the terms "justify" and "guarantee". Lizbeth decided to keep her initial meaning for "justify" (her assumption bAL1) during Discussion 1 and, instead, she employed the new terminology introduced during the intervention (e.g., "guarantee", "sufficient") to make precisions about the status of confirming examples when proving UAS.

Establishing meanings for "justify" and "sufficient to guarantee"

Both Discussions 1.4.1 and 1.4.2 used the same classroom episode as a frame of reference[109]. The context consisted of a third-grade class that was asked to analyze the truth value of the statement St16 *"All divisions of natural numbers are exact divisions"* Students provided confirming and contradicting examples as part of their analysis. The discussions had a focus on two different issues. While Discussion 1.4.1 asked whether the confirming example *"5 divided by 1"* justified that St16 was false, Discussion 1.4.2 asked whether the same example justified that St16 was true. The teachers had already disproved St16 during Discussion 1.0. Lizbeth initially considered that the confirming example *"5 divided by 1"* disproved St16; however, she changed her answer as she realized that it was *"not a counterexample"*, but instead it was *"affirming that it is exact"* (turn 10 in Episode 7 above, see Section I.2.2.1) and as such it could not justify that the statement was false (*confirming examples do not disprove false UASs*). Lizbeth's answer was reinforced by Gessenia's observation that the example in discussion in fact *supported* the statement (turn 13 in Episode 7). This might have contributed to strength Lizbeth's initial use of "justify" as support, verify and confirm and create the idea that even though the statement was clearly false, a confirming example could still "justify" that it was true. This became clearer afterwards, during Discussion 1.4.2.

In Discussion 1.4.2 the teachers were asked whether the (confirming) example *"5 divided by 1"* justified that St16 was true. Lizbeth, like Gessenia but unlike Andrea, answered affirmatively. Episode 14 includes the discussion developed about this issue, where Lizbeth's use of "justify" was more explicit.

[109] See Appendix D1 for details on the content of these discussions.

Episode 14

Turn	Who	What
1	Lizbeth	*I say yes, because with that [example] he has confirmed that his statement is true.*
2	I	*The question is, is it sufficient to justify that the statement is true?*
3	Andrea	*No.*
4	I	*Why not?*
5	Andrea	*Because not always only one example is going to be necessary [most likely she means "sufficient"], but in this case more examples are needed. If I give more examples-,*
6	I	*How many more examples?*
7	Andrea	*No, that's why, it depends. What I have to look for-, if I'm given a statement, I have to look for a counterexample.*
8	I	*What you [Gessenia] said before [in the previous discussion], you said, this is an example that supports the statement. And it supports the statement, doesn't it?*
9	Gessenia	*Sure, it supports the statement.*
10	I	*It supports the statement in the sense that, WELL, it SEEMS that this statement is true. But, because of this example, am I going to assert that this is true? In other words, is it going to be sufficient that with this example I conclude that, thanks to this example, my statement is true?*
11	All teachers	*No.*
12	Lizbeth	*But, what happens if he looks for another example that still confirms that this is true? For example, 20 divided by 5.*
13	I	*Ok then, 20 divided by 5. Let's suppose that the class says no, but teacher, 20 divided by 5, 5 divided by 1, 18 divided by 3.*
14	Lizbeth	*12 divided by 4*
15	I	*What would you do? You can have this situation in which the students start to provide examples that support the statement. What would you do?*
16	Andrea	*I show my counterexample and prove that not all.*

Lizbeth explained why she claimed that the confirming example *"5 divided by 1"* justified that the (false) universal statement St16 was true (*"I say yes, because with that [example] he has confirmed that his statement is true"*, turn 1). Her explanation is clear evidence of her interpretation of "justify" as confirm, verify, support. Her use of "justify" might be linked to its usage in scientific disciplines different from mathematics (see Reid & Knipping, 2010; Krantz, 2011). In other scientific fields, like in biology or physics, experimentations and generalizations are at the core. Scientists provide convincing evidence to prove something. That evidence, though, is different in nature from mathematical evidence. Lizbeth's initial use of "justify" was in that sense similar to that deployed in other (non-mathematical) fields. Beyond that, Lizbeth assumed that it was acceptable that even though a statement was false, confirming examples *justified* that it was true because of their supporting nature (they support or confirm the statement in those cases). For Lizbeth there was no room for a conflict between her awareness that the statement was false (she herself had disproved it before) and her personal assumption that a confirming example "justified" that the statement was true. The latter was based on the meaning she used for "justify" at that moment ("justify" means confirm, support, verify). It is not that she considered that the statement was true[110], but she was aware that there

[110] A similar episode arose after the intervention for teachers, during Lizbeth's teaching of her Session 12. Her class analyzed a false UAS that allowed the possibility of supporting examples, for which Lizbeth categorically concluded that its falseness was guaranteed by the existence of a counterexample.

were examples that confirmed the statement (turns 12 and 14). It is in that sense that both of her assertions could co-exist without any apparent contradiction.

Notice that when the task was rephrased (turn 10), Lizbeth acknowledged that the confirming example was *not sufficient to conclude* that the statement was true (turn 11). This is interesting as it already showed that Lizbeth used the verb "justify" to mean two different things in the context of discussing a false UAS that admitted confirming examples: (1) "justify" means *is sufficient to conclude* (e.g., 15 justifies that *"All numbers divisible by 5 are even"* is false, and (2) "justify" means *verify, support, confirm* (e.g., 20 justifies that *"All numbers divisible by 5 are even"* is true). Lizbeth was aware that a counterexample refuted a false UAS or, in other words, a counterexample *justified* that the statement was false. The usage of "justify" in this context is as in the former meaning; that is, a counterexample is sufficient to conclude that a statement is false. On the other hand, Lizbeth used "justify" to claim that confirming examples supported, verified that the statement was true, which is the second meaning.

Lizbeth used again her dual interpretation of "justify" during later discussions and for other cases of false UASs. For example, during a Recap for Discussion 1 the teachers were asked to provide examples of false universal statements. At that moment we were focused on Andrea's statement St42:

St42: *All numbers divisible by 3 are even numbers.*

Gessenia suggested 15 as a counterexample that disproved St42. Lizbeth agreed with it. Then attention was drawn to the set of analysis and the number 6 was put forward to determine whether it justified that the statement was true. Lizbeth answered affirmatively. She claimed that number 6 "justified" that the statement was true, even though she was aware that the statement was false. The truth value of the statement was brought up and Lizbeth explicitly stated: *"but I must keep looking"*, presumably for a counterexample. This suggested that, according to Lizbeth, even though number 6 *justified* that the statement was true (in the sense of "verified" it), this confirming example alone was insufficient to conclude that the statement was true. This revealed again Lizbeth's use of "justify" as confirm/verify, but also her awareness that a confirming example did not suffice to conclude that the statement was true.

This episode showed that Lizbeth was not thinking about UASs in three related ways that have been reported in the literature. She did not assume that the false UASs with possible confirming examples (e.g., St42) we analyzed could have been *both* true and false (see e.g., Buchbinder & Zaslavsky, 2009), nor was she unsure about the truth value of the statements (see e.g., Barkai et al., 2002), nor did she consider that a counterexample was an exception for the statements (see e.g., Reid, 2002).

Agreeing on meanings

In a subsequent discussion the teachers explained the way they understood the expression "valid justification". Through her explanation, Lizbeth made explicit again the use of her meaning for "justification" as well as her contrast with her meaning for "sufficient to guarantee".

During Discussion 3.2[111] Lizbeth referred to "confirmation" as a key aspect of a "valid justification" (*"He is claiming, he is confirming that what he has done is true"*, turn 3 in Episode 12 above). Hence, Lizbeth used both "justification" and "valid justification" with

[111] For more details about the task in Discussion 3.2, see the context described before Episode 12 (above).

the same meaning, namely: confirmation, verification. Furthermore, Lizbeth was aware that the three confirming examples Pepito provided did not guarantee or were not sufficient to show that his conjecture (St72: *All palindrome numbers are divisible by 11*) was true since, as she put it, *"he [Pepito] did not analyze all palindrome numbers"* (turn 7 in Episode 12). Implicitly, according to Lizbeth's words, the confirmation of all the cases involved in a universal-statement conjecture *sufficed to guarantee* that the conjecture was true (her assumption dAL3[1]).

In short, Lizbeth showed the following current assumptions about confirming examples and their status when proving UASs: (1) confirming examples *justify* that a UAS is true (whether it is true or false); (2) in order to *guarantee* that a UAS is true, verifying a few (and not all the) examples involved in the conjecture is not *sufficient*, verifying all the cases involved in the statement is a must. While her first assumption is still a manifestation of her initial meaning for "justify" (assumption bAL1), the second is a refinement she built during the intervention and, notably, from the logical interpretation of the UAS we have analyzed. This suggested that Lizbeth accommodated the meaning she initially used for "justifications" for true UASs in order to extend it to the case of false UASs to mean *confirm/verify/support*. As her vocabulary expanded during the intervention, by adding expressions like "sufficient to guarantee", and the logical interpretation of universal statements was put forward, she separated the two processes – "justifying" and "guaranteeing".

Part of our discussion seems to have triggered Lizbeth's reconsideration of her distinction for the meanings of "prove" and "sufficient to prove or guarantee". With one of my inputs I underscored that a "valid justification" was a justification that proved that the statement was true (*"When I ask if a justification is valid, I am asking whether the justification indeed PROVES that the statement is true, since that is what Pepito claims in this case."*, turn 4 in Episode 12). Andrea added that it meant that the justification was *sufficient* (turn 5), which Lizbeth might have found against her own meanings since she distinguished between a "justification" and a "sufficient justification to guarantee something". My explicit agreement with Andrea's remark, plus the explicit link I made with the term "guarantee" (turn 6) might have pushed Lizbeth into re-accommodating her personal assumptions for the terminology she used. Her new assumption was evidenced at the end of Discussion 3.2, where the teachers shared, as a recap, the things they had learned so far from Discussion 3. In particular, Lizbeth pointed out again, though in general terms, the need to analyze each case involved in the statement in order to *justify* that it was true (her assumption dAL3[2]). The gestures she used during her explanation suggested her focus on the logical interpretation of the statement.

> Lizbeth: *That we should analyze, according to the statement we have, we should analyze, from all that is indicated, analyze every one (Lizbeth uses gestures with her hands to suggest that she means each case), every one, in order to assert that it is true, in order to justify it.*

Here she did not claim that the verification of some confirming examples would prove/justify that the statement was true, as she had previously done; in contrast, she emphasized the need to analyze every case involved in the statement in order to justify it. This means that at this point Lizbeth began to apply her meaning for "sufficient to guarantee" to "justify". Presumably, this was a result of explicitly discussing the meanings we independently used for "valid justification" and explicitly establishing a common meaning to be used, during Discussion 3.2.

Although Lizbeth's ways to express her reasoning during the intervention seemed puzzling or contradictory, she managed to come to understand that when proving a UAS that involved an infinite number of cases, verifying a subset of confirming examples did not *suffice*. Instead, as she put it, all the cases involved in the statement should be verified, even though in many cases she struggled to show this. That means that the development of Lizbeth's understanding about proving UASs revealed two levels of awareness she gained from the intervention. On one hand, she was aware that some examples might confirm or verify an infinite US in the sense that the statement is particularly true for those cases; on the other hand, she was aware that verifying confirming examples was not *sufficient to guarantee* that an infinite US was true.

An imaginary UAS

The analysis of an imaginary UAS revealed the challenge that Lizbeth faced when identifying the characteristics for a sufficient justification to prove it. Such a challenge apparently stemmed from the task formulation combined with her interpretation of the word "suppose".

During Meeting #10, after the intervention, the teachers solved the Extra Activity 1, which was expected to be solved individually first and discussed in a whole-group discussion afterwards. Task 1 consisted of the *imaginary* statement St139,

> St139: *All natural numbers bigger than 5 are RAINBOW numbers*.

Specifically, Task 1.1 asked the teachers to suppose that St139 was true and they were asked to answer what kind of mathematical evidence would suffice to show in order to guarantee that it is true. Lizbeth's first response was *"an example that confirms what is claimed in the statement"* is sufficient to show that the statement is true. During the whole-group discussion Lizbeth exhibited her struggle with the first part of the task: "Let's suppose that [St139] is a true statement".

> Lizbeth: *But here that is asserted, suppose that this [St139] is true. I mean, this is like asserting, isn't it?*
>
> I: *Here you are asked for the type of evidence that guarantees that.*
>
> Lizbeth: *Verify every number.*
>
> I: *How many numbers?*
>
> All teachers: *Infinite.*
>
> Lizbeth: *Then it is not sufficient, but use the numbers-, here I got confused. I mean, if it states that this is true-, then it does not state that this is true, is this [the truth value of the statement] unknown yet?*

In the excerpt Lizbeth showed that after the intervention she was still aware that every case involved in the statement needed to be verified in order to guarantee that it was true (her assumption dAL3[2]); however, her confusion (as she put it) stemmed from the assumption the teachers were requested to make (Let's suppose that [St139] is a true statement). In fact, this request was included as a hypothetical situation, to avoid that the teachers might refuse to deal with the truth value of St139 because it was an imaginary statement for which a truth value did not exist. Lizbeth understood the task as if the statement was indeed true and they needed to show an example that confirmed this.

Unlike the other statements for which Lizbeth developed her new assumptions about proving UASs, St139 was imaginary and thus its truth value was impossible to be determine. The focus on an imaginary statement increased the difficulty of the task. In particular, the request to suppose that the statement was true misled Lizbeth and she gave an answer that was not exactly aligned with her current awareness of what was entailed in proving an infinite UAS.

Figure 26 shows the development of Lizbeth's assumptions about confirming examples and their status when proving UASs and the elements or aspects of the intervention that might have supported or triggered changes[112].

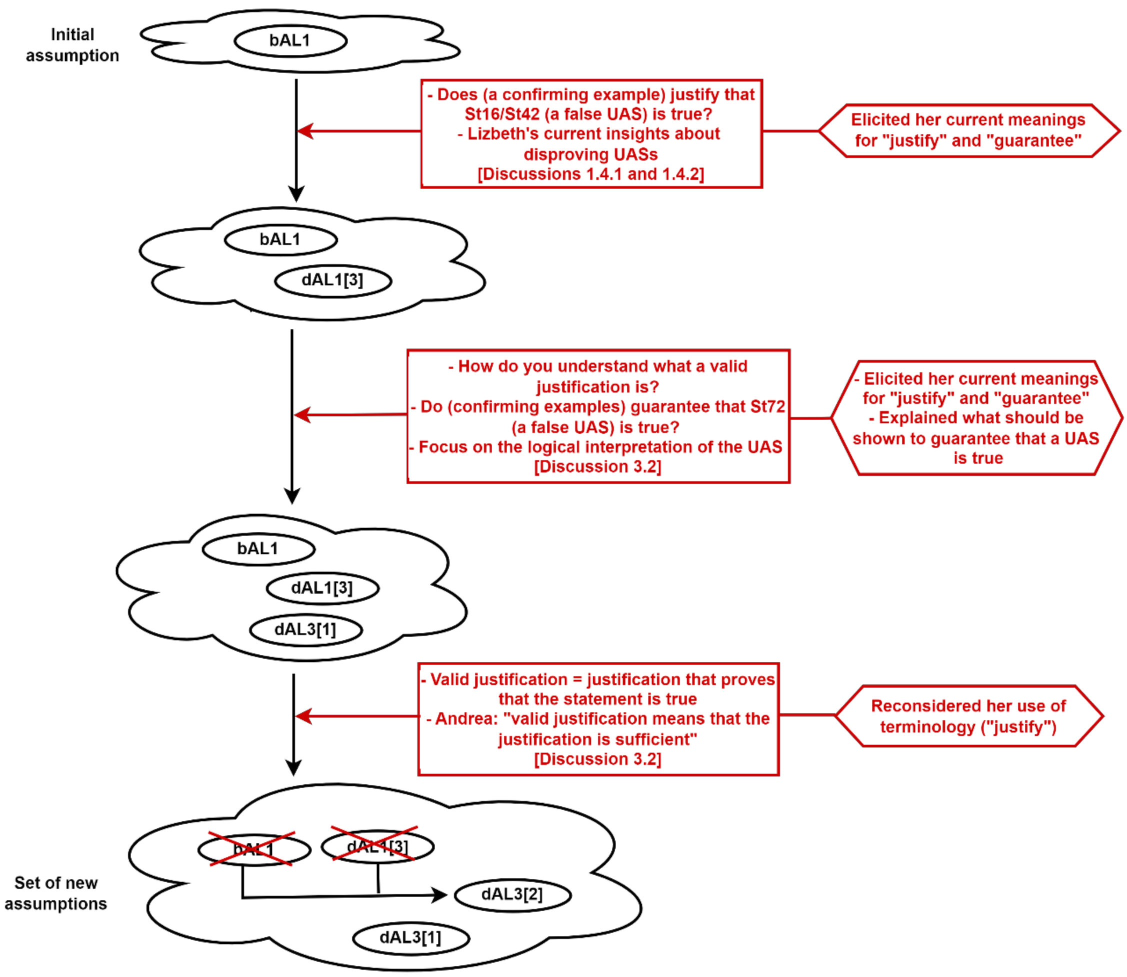


Figure 26. The development of Lizbeth's assumptions about confirming examples and their status when proving UASs (bAL1: Confirming examples can justify that a UAS that involves infinite cases is true; dAL1[3]: Confirming examples prove/justify a false UAS that admits confirming examples with infinite cases involved is true, though they do not necessarily guarantee that the UAS is true; dAL3[1]: Verifying some examples does not guarantee that a UAS is true. All examples need to be verified in order to do that; dAL3[2]: In order to justify that a statement is true, we should verify each case involved in the statement)

---

[112] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

Summary of Section I.3.1

Before the intervention the three teachers shared a very common assumption for how to prove a true universal statement. The teachers used and accepted confirming examples to prove infinite universal statements. However, there were important differences in the approaches they took. While Andrea's attention was on showing confirming examples that exhibited a *pattern* in order to focus on generality, Gessenia and Lizbeth used random examples that verified the US, without any attempt to focus on general aspects, which exhibited Gessenia's and Lizbeth's naïve empiricism proof scheme.

The development of Andrea's assumptions about the status of confirming examples when proving USs was gradual and experienced several refinements. Her process of understanding shows that her attention to the logical interpretation of USs and the number of cases involved during the intervention played a crucial role. Her first change of assumption was extreme as she assumed that confirming examples could not prove a statement, which she later specified meant a universal statement. This first change of her initial assumption was presumably provoked by: (a) the evaluation of arguments during the first part of the intervention, from which Andrea seems to have overgeneralized certain features (e.g., USs are proved with general text-form properties or definitions), (b) the incomplete inputs I provided during the first part of the intervention (e.g., a verification is a confirmation that the statement holds with the use of examples, whereas a mathematical justification goes beyond a verification as it might involve complex ideas), and (c) the focus on false USs and their disproving during Discussion 1 of the second part of the intervention (Andrea assumed that examples were valid only when disproving USs). By challenging her new assumption, Andrea brought up the "size" of the set of analysis. Andrea came up with a new assumption that had a focus on the case of finite USs, which can indeed be proved by confirming examples (a verification of all the finite cases involved). The case of generic examples opened new possibilities when proving infinite USs, which Andrea rejected at the beginning. A discussion on what other considerations needed to be addressed when generic examples were used in valid proofs permitted that Andrea accepted them as valid proofs.

On the other hand, an interesting phenomenon emerged with the development of Lizbeth's assumptions. It was related to her interpretations of expressions such as "justify" and "sufficient to guarantee". Lizbeth initially used the verb "justify" to mean verify, confirm, support, which was consistent with her initial approach to prove/justify true USs. During the first proof-related discussions Lizbeth also used that meaning for the case of false USs that allowed confirming examples. Lizbeth made claims where she accepted that confirming examples justified that those kinds of USs were true. During the intervention new expressions were introduced (e.g., "sufficient to guarantee", "valid justification"). This, together with her new insights about the logical interpretation of a US and disproving USs, supported Lizbeth's refinement of her previous assumptions that entailed a distinction she made between "justify" and "sufficient to guarantee". Lizbeth claimed that confirming examples could justify (in the sense of *verify*) that a false US that admitted confirming examples was true (for those concrete examples). Yet, she clarified that those confirming examples were *not sufficient to guarantee* that the US was true. Learning about the logical interpretation of USs helped Lizbeth to move from a naïve empirical proof scheme to a general approach that focused on the verification of all cases involved in the statement. Lizbeth's meaning for both expressions ("justify" and "sufficient to guarantee") merged as a result of a discussion we had during the intervention where the teachers shared their interpretations for "valid justifications" and we agreed on a common meaning.

Even though Lizbeth's assumptions about the status of confirming examples when proving UASs evolved according to the mathematical standards of proof, she still struggled to prove UASs. This difficulty was manifested through her recurrent reliance on confirming examples despite of her awareness about the insufficiency of those verifications to prove a true UAS (as I also show in the next section). On the other hand, Andrea was more cautious about this matter and openly acknowledged that she could not find a sufficient justification for a true universal statement when that happened. The teachers' lack of success when proving true USs can be explained in terms of their lack of familiarization with proof methods different from direct proofs.

3.2. Choice and use of examples when formulating, evaluating and attempting to prove a US conjecture

In the previous section I showed the development of Lizbeth's assumptions related to the status of confirming examples when proving UASs. In this section I focus on her choices and uses of confirming examples when formulating, evaluating and attempting to prove a false universal-statement conjecture.

Lizbeth revealed her criteria to select examples when solving a task after Discussion 3.2 and afterwards, during a recap for the previous immediate discussions. After Discussion 3.2, the teachers were asked to solve a task that consisted in identifying a pattern, formulating a conjecture and evaluating the truth value of the conjecture. Specifically, they were given the expression $1 + 1141n^2$ and asked to answer whether such expression produced a perfect square number for $n = 1; 2; 3$. Then they were requested to share their observations, to formulate a conjecture if possible, and finally to evaluate whether their conjecture was true and explain why[113]. I expected that the teachers become aware and convinced that the only way to guarantee that a conjecture is true is by proving it. They were expected to notice that proving a US conjecture involved showing the general reasons for why it was true for all the cases involved in the statement. This implied that no big random confirming example was sufficient to show that, especially considering that the task involved a false conjecture with an "extreme counterexample"[114].

In the following I go into details about the criteria that Lizbeth employed for her choice and use of examples depending on the activity she paid attention to: Increasing ordered sequence of examples to formulate a conjecture; examples that increased in complexity and organized by properties to evaluate and attempt to prove a conjecture; random examples to support the search for counterexamples.

Lizbeth's choice of increasing ordered sequence of examples to formulate a conjecture

Lizbeth's selection of examples was directly linked to the activity she was focused on. In order to identify and formulate a conjecture that encompassed all non-zero natural numbers, Lizbeth concentrated on testing the first ten non-zero natural numbers, one by one in increasing order. The process of formulating the conjecture revealed Lizbeth's struggle to express her ideas. Her first formulation was:

[113] See Appendix D3 for the task content.

[114] "Extreme counterexample" in the sense that it is a very large number (a 26-digit number) that was found to be a counterexample by using computers. The task was adapted from G. J. Stylianides and A. J. Stylianides (2014).

> Lizbeth: *For the numbers 1, 2 and 3 it [the formula $1 + 1141n^2$] does not produce a perfect square… Now, I have a conjecture. It says, when a number, a natural number, multiplied by n squared, which could be any number, plus one, it won't have a whole square root, it won't have a perfect square root. It won't have square root. Could it be?*

Her second attempt to formulate the conjecture, which entailed her substitution of the following seven natural numbers, was more comprehensible.

> Lizbeth: *When a natural number [different from zero] is substituted in $1 + 1141n^2$, it [the result] will not be a perfect square number.*

Her choice for the first ten non-zero natural numbers led her to coming up with the expected conjecture. While Lizbeth's use of the three first non-zero natural numbers already supported her *realization of the pattern*, the subsequent seven natural numbers supported a *refinement* of her initial formulation of the conjecture[115].

Despite the persuasive nature of the first ten non-zero natural numbers, Lizbeth was certain that those ten verifications did not guarantee that the conjecture was true, which was consistent with her assumption dAL3[1] (see Section I.3.1.2 above). This means that according to Lizbeth those ten confirming examples were insufficient to prove the conjecture. Verifying the first ten non-zero natural numbers might have not increased Lizbeth's confidence on the truth of the conjecture, but it helped her see what the pattern was and express her conjecture more clearly.

Lizbeth's choice of examples that increased in complexity and organized by properties to evaluate and attempt to prove a conjecture

The process of evaluating the conjecture revealed Lizbeth's choice for examples that increased in complexity and that at the same time were organized by "properties" or "categories"; that is, examples clustered by specific features, as if she tried to break down the set of natural numbers into "classes". All categories shared the same feature of being confirming examples, which increased Lizbeth's confidence in the truth of the conjecture. This led Lizbeth to attempt to prove the conjecture, unsuccessfully.

The criteria Lizbeth used to group the examples she tested was to consider, first, numbers smaller than 100. Besides the first ten non-zero natural numbers she verified, she chose $n = 22$ and 66 and concluded that *"We have already seen that with numbers smaller than 100, it is not [a perfect square]"*. Second, she chose numbers bigger than 100. She verified the conjecture for $n = 120$, 1555, 9999, 200 and 111, in that order. Presumably, those categories were intended to be complementary[116], as if she had expected to cover the complete spectrum of the non-zero natural numbers (the set of analysis). She explained that her breaking out of her "increasing-pattern" after she tried 9999 was based on two features of 200: its parity as well as it ending in two zeros, which underscored a more specific property inside such category.

> Lizbeth: *until now what I've been doing, let's say, hmm, I wanted to try, well, I linked it with the fact that 200 is an even number… That's why. Then I said let's*

[115] The second version of the conjecture was an improvement of the first one, which at first Lizbeth did not find easy to clearly put into words.
[116] Even though Lizbeth forgot to include number 100 in one of the two groups.

> *see, I'm going to see, because I've been looking for-, sure, other ones were also even, but I said, this is a number, let's say, that ends with zeros.*

The reason for her use of examples with repeated digits, like 1555, 9999 and 111, was revealed later, when she showed her already increasing conviction that the conjecture was true (*"I believe that yes [it is true], because when taking random numbers*[117]*-, right? And we see that suddenly taking equal digits, others ending in 0"*). Lizbeth's intention to encompass all non-zero natural numbers (due to her awareness that in order to prove the conjecture, she needed to verify all cases involved in the statement[118]) was shown through her failed attempt to prove the conjecture "locally" in subsets of the set of analysis.

It was interesting to see the development of Lizbeth's reasoning. She organized examples according to properties to try to conclude that the conjecture was true. Her approach, though unsuccessful, attempted to address the issue of generality since Lizbeth was aware that the statement was about an infinite number of cases (all non-zero natural numbers). Unlike her initial naïve empirical approach to prove universal statements[119], here Lizbeth used a strategy to choose her examples. In her process of evaluating the truth value of the conjecture Lizbeth revealed her engagement in a form of reasoning she had not shown so far. Lizbeth's emergent approach to aim for generality was new to her. She had not received any previous instruction about or engaged in proof-related activities in mathematics[120].

According to Lizbeth, the first sixteen examples she had tested, though grouped by categories that were intended to encompass all the cases involved in the statement, did not *guarantee* that the conjecture was a mathematical truth. Despite that, she acknowledged more than once that she *believed* that the conjecture was true. This means that even though she had the expectation that the conjecture was true, she was aware that verifying those examples was insufficient to prove that the conjecture was indeed true. This revealed again the status she assigned to the verification of a subset of the infinite set of cases involved in the statement; namely, they are insufficient to prove a universal-statement conjecture.

Lizbeth's selection of examples that increased in complexity involved her use of numbers of one, two, three and four digits (*"And bigger every time and it [the conjecture] still holds"*). Likewise, her choice for examples based on properties included her selection of numbers smaller than 100 and numbers bigger than 100. With her strategies Lizbeth intended to cover the whole set of cases involved in the conjecture, which implied that she attempted to focus on generality even though she still lacked the methodological tools to achieve that goal. Lizbeth's strategy resembles some of the approaches used by mathematicians to choose examples when engaged in proof-related activities that were reported by Lockwood et al. (2016). Moreover, Lizbeth showed that she was aware that the examples she used, which verified the conjecture, were insufficient to prove the conjecture. Lizbeth's awareness of the activity with examples and the way it relates to the broader proof activity is one important aspect Lockwood et al. (2016) emphasized in the work of mathematicians. It was interesting to see that at least at a first stage Lizbeth began to develop that kind of awareness and that it influenced her own regulation of her choice and use of examples. This improved form of reasoning was, in my view, a result of the

[117] Lizbeth called her choices "random examples"; however, they shared common characteristics, like being numbers with equal digits, etc. See the examples she chose.
[118] For details, see previous Section I.3.1.2.
[119] See her initial assumption bAL1 in Section I.3.1.2.
[120] She said so during the first part of the First Exploratory Interview, before the intervention.

discussions developed about the status of confirming examples when proving UAS is involved (see Section I.3.1 above). Based on such awareness she could move forward and focus on generality as an important step towards the actual proving of universal statements.

Lizbeth's use of random examples

Another important use Lizbeth gave to examples when evaluating a universal statement had to do with the role she attributed to random examples. During a recap for the task after Discussion 3.2, Andrea put forward the status of random examples when proving universal statements. She made the claim that random examples did not count to prove universal statements. In contrast, Lizbeth disagreed with Andrea's claim. Episode 15 includes the dialogue developed about this debate.

Episode 15

Turn	Who	What
1	Andrea	*[Checking] random cases do not count [to prove].*
2	Lizbeth	*Not always. Random cases could count. Well, the conjectures-, don't you say that unless we prove it, or we have the justification, they are not mathematical truths?*
3	Andrea	*Sure.*
4	Lizbeth	*So, if you give a random one and you find a justification, it will be a mathematical truth.*
5	Andrea	*No, in order to say that it is true, one, one-*
6	Lizbeth	*We have to justify it.*
7	Andrea	*No matter if they are random, they do not count to justify.*
8	Lizbeth	*But do you realize that they [random examples] do not count if you do not find a counterexample?*
9	I	*(I repeated what Lizbeth said)*
10	Andrea	*Never mind. Random cases are not valid.*
11	I	*Let's see. The random examples you [Lizbeth] have found, do they count to conclude that this conjecture is true?*
12	Lizbeth	*No.*
13	I	*Then, in what sense do you assert that they count? That is what we are trying to understand.*
14	Lizbeth	*Ah, alright. No, but I was talking about, let's say, the conjectures that could be given, right? You cannot assert a conjecture is true or false until you find its justification.*

Lizbeth and Andrea's disagreement seems to stem from the different foci of their claims. While Andrea's claim was centered on the *insufficiency* of random examples *to prove* a universal statement, Lizbeth's claims were focused on the *role* that random examples might play when *evaluating* and even when *attempting to prove* a universal-statement conjecture.

Presumably, Lizbeth tried to explain that random examples were important when determining whether a conjecture was true or not for two reasons: First, she seems focused on the importance of random examples when trying to find a justification that *proves* that a universal statement is true (*"So, if you give a random one and you find a justification, it will be a mathematical truth"*, turn 4). Recall that at this point of the intervention Andrea was aware that in order to *justify* that a statement was true, each case involved in the statement should be verified[121]. Notice that Lizbeth chose "random" examples strategically in her attempt to cover all the set of analysis (*"I believe that yes [it is true], because when taking random numbers[122]-, right? And we see that suddenly taking equal digits, others ending in 0"*). It is possible that she sees her choice of random examples as a chance to reveal the general reasons for why the statement is true. Second, Lizbeth seems aware that this process might support finding a counterexample (*"But do you realize that they [random examples] do not count if you do not find a counterexample?"*, turn 8), which would lead to falsifying the universal statement[123].

Considering the complexity of the task that the teachers previously discussed, turn 8 (*"But do you realize that they [random examples] do not count if you do not find a counterexample?"*) also suggests that Lizbeth was cautious about the possibility of evaluating similar US conjectures with counterexamples that were hard to find. Lizbeth seems to refer back to the conjecture they had analyzed and the extreme unexpected counterexample it had. In this case, Lizbeth seems to have assumed that random examples should not be undervalued since one of them could turn to be a counterexample.

The case of Lizbeth is similar to those previously reported in the literature where individuals struggle to find a counterexample, which might push them to assume that a (false) US is true (e.g., Ko & Knuth, 2009a)[124]. Here Lizbeth *believed* that the statement was true, because of the several strategically chosen confirming examples she verified and the approach she took to try to prove the conjecture; however, her insights about what was involved in proving a US prevented her from making such a strong conclusion.

Summary of Section I.3.2

Lizbeth's insights about what is involved in proving USs (see Section I.3.1.2 above) influenced her choice and use of examples as she engaged in formulating, evaluating and attempting to prove US conjectures. She began the intervention with a very naïve selection of easy-to-compute examples in order to show that a US was true. As she understood that normally a few confirming examples did not suffice to guarantee that a US was true[125], she focused on a systematic criterion to choose examples with different purposes. This was revealed through her solving process of a task that involved a false US with a hard-to-find counterexample.

Her selection of examples was directly linked to the activity on which Lizbeth intended to focus. For example, Lizbeth verified one by one the first ten non-zero natural numbers to find a pattern that would lead her to formulating a conjecture that encompassed all non-zero natural numbers. Lizbeth's evaluation of the conjecture involved her choice of random examples. For Lizbeth, random examples might have resulted in counterexamples that would refute the conjecture, or in confirming examples that suggest the general reasons for why the conjecture was true. The confirming nature of the examples she tested increased Lizbeth's confidence about the "trueness" of the conjecture. Lizbeth's (failed) attempts to prove the conjecture included a systematic selection of examples. She

[121] See Lizbeth's assumption dAL3[2] in Section I.3.1.2 above.
[122] Lizbeth called her choices "random examples"; however, they shared common characteristics, like being numbers with equal digits, etc. See the examples she chose.
[123] See Section I.2.2.1 for details on Lizbeth's assumptions related to this issue.
[124] For details, see Chapter 2, Section I.5.3.
[125] This is Lizbeth's assumption dAL3[1] in Section I.3.1.2.

organized her examples by “properties”; that is, examples with common characteristics like, numbers ending in zeros or numbers with equal digits.

Even though the confirming nature of the examples she chose in this process strongly contributed to Lizbeth’s belief that the conjecture was true, she was aware that her procedure ultimately was insufficient to guarantee that the conjecture was a mathematical truth. Her choice of examples by properties was aimed at gaining insight into proving the conjecture. She attributed representativeness to those examples, in a way similar to dividing the set of natural numbers into “classes”. Her approach revealed Lizbeth’s intention to entail all on-zero natural numbers by proving the conjecture locally within the “classes”. With it, Lizbeth attempted to address generality, which began to be part of her attention during the intervention.

3.3. Emergent assumption about proving UASs as eliminating the possibility of counterexamples

Lizbeth’s understanding of proving universal statements experienced some changes. Initially she assumed that verifying a few examples sufficed to prove a universal statement. Later she became aware that those verifications were not sufficient to *guarantee* that a universal statement was true[126]. At one point of the intervention it was interesting that Lizbeth drew attention to an additional criterion to prove universal statements: the non-existence of counterexamples. In this section I focus on the development of this assumption.

In order to give some context, I first include some previous relevant events that are linked to the moment when I observed Lizbeth’s emergent assumption. Discussions 1 and 2 were mainly focused on false UASs and disproving them[127]. Based on those discussions the teachers became aware that a single counterexample was sufficient to disprove a universal statement. On the other hand, Discussions 3.1 and 3.2 had a focus on explicitly reflecting about the link among “conjecture”, “justification” and “mathematical truth”. Specifically, Discussion 3.2 included the case of a (false) conjecture formulated by an imaginary student (Pepito), who explained that his conjecture was true because of his verification of four “big” supporting examples. The main goal of the discussion was to engage the teachers in arguing whether they regarded Pepito’s answer as a valid or invalid justification and why.

During a Recap for Discussions 3.1 and 3.2 the teachers were requested to share the insights they gained from those previous discussions.

> Lizbeth: *That we should analyze, according to the statement we have, we should analyze, from all that is indicated, analyze every one (Lizbeth uses gestures to suggest that she means each case), every one, in order to claim that it is true, in order to justify it.*
>
> I: *And what about when we make generalizations, when we make conjectures.*
>
> Gessenia: *That it is not true that when a pattern repeats it is going to be necessarily justifiable.*

[126] For details in this development, see Section I.3.1.2 above.
[127] For details, see Section I.2 above.

> I: *(I provided an example of a false generalization: "all men are liars") I can always make generalizations, can't I? That is different from concluding that this is indeed true, isn't it? In order to give that additional step…*
>
> Lizbeth: *I have to justify… Just then is when this is going to become a mathematical truth-, no, depending on the justification.*
>
> I: *Why depending on the justification?*
>
> Gessenia: *But if I already have the justification, I mean, if I already have the justification, it means, it is a [mathematical] truth.*
>
> Lizbeth: *But then I would add there, if we do not find any counterexamples, then it [the conjecture] would be a [mathematical] truth.*

The way Lizbeth stated her assumption dAL3[3] (*"if we do not find any counterexamples, then it [the conjecture] would be a [mathematical] truth"*) may be interpreted at least in two different ways. (1) Finding no counterexamples, perhaps because they are not accessible to the prover[128], even though counterexamples may exist, might mislead the prover to rush and conclude that the conjecture is true. (2) Finding no counterexamples, because the possibility of the existence of any counterexample is eliminated, that is what guarantees that a conjecture is certainly true. I believe that Lizbeth's assumption was aligned with the second interpretation. My presumption is based on the following rationale: first, the teachers had become aware that while there was still a chance for counterexamples to exist, a conjecture could not be guaranteed to be true. For example, during Discussion 3.2 the teachers analyzed the example of a conjecture suggested to be true based on four (some of them "big") confirming examples that verified the conjecture; however, the conjecture turned out to be false, as the teachers found counterexamples that refuted it. Second, in the first insight that Lizbeth shared, she referred to the need to verify each case involved in a universal statement in order to prove it (*"That we should analyze, according to the statement we have, we should analyze, from all that is indicated, analyze every one, every one, in order to claim that it is true, in order to justify it"*). This suggests that she implicitly considered every element in the set of analysis and did not conceive the idea that a counterexample could escape that control. Third, she was aware that the existence of at least one counterexample would refute a universal statement[129]. Hence, being aware of the possibility of missing one counterexample does not seem a plausible interpretation. Therefore, eliminating the possibility to find counterexamples for a US conjecture seems to have become relevant for Lizbeth, as a guarantee that the conjecture is indeed a mathematical truth.

4. Negation of Universal Affirmative Statements

In this section I focus on the development of Andrea's assumptions about the negation of universal affirmative statements. Before I get into details, I introduce specific language to distinguish forms of negation that I found to be relevant during my data analysis.

It is important to understand that a negation can be expressed in different ways. I call *implicit negations* those negations that are already included in the statement itself (e.g., *"**Not** all puppies are mischievous"* or *"**It is not the case that** some even numbers are palindrome numbers"*). In contrast, I call *explicit negations* those negations that are an

[128] See for example the case in Section I.3.2, above.
[129] For details on this issue, see Section I.2.2.1 above.

explicit request to negate a statement (e.g., "**Negate** the statement *'All puppies are mischievous'*").

Implicit negations may exist in a simple form, but not all implicit negations have a simple form (see Table 14). For example, "all-statements" and "there-are/exist-statements" both admit simple implicit negations (e.g., *"**Not** all X are Y"* and *"There **are no/does not exist** X that is Y"*, respectively). In contrast, "some-statements" and "if-then-statements" do not allow these simple forms. Implicit negations of "some-statements" look wordier and more complex (e.g., *"**It is not true that** some X are Y"*). That is, they involve using expressions such as *"it is not the case"*, *"it is not true"*, *"it is false that"* to the front of the main statement.

Table 14. Negations and their forms of expression.

		Forms of expression for negations		
		Implicit Negation *(It is not the case, it is not true, it is false that...)*	**Simple Implicit Negation**	**Explicit Negation** *(Negation of...)*
Universal Statements	*"All X are Y"*	*It is not the case that all X are Y*	*Not all X are Y*	*Negation of "All X are Y"*
	"If X, then Y"	*It is false that if X, then Y*	-------	*Negation of "If X, then Y"*
	"No X is Y"	*It is not true that no X is Y*	-------	*Negation of "No X is Y"*
Existential Statements	*"Some X are Y"*	*It is not the case that some X are Y*	-------	*Negation of "Some X are Y"*
	"There exists X that is Y"	*It is not true that there exits X that is Y*	*There does not exist X that is Y*	*Negation of "There exists X that is Y"*
	"There are X that are not Y"	*It is false that there are X that are Y*	*There are no X that is not Y*	*Negation of "There are X that are Y"*

Andrea's assumptions about negation of UASs include her initial assumptions about the relation between negation and falsity, and equivalences of the *simple implicit negation* of a UAS. I describe the influence of everyday language in Andrea's initial approach to negating UASs and I discuss the factors that led her to change her approach.

Table 15 includes Andrea's assumptions related to the negation of universal affirmative statements.

Table 15. Andrea's assumptions about the negation of UASs

Assumption	**Type of assumption (when first observed)**
dAA1[4]: *"Not all X are Y"* and *"It is false that all X are Y"* state the same	Initial assumption (during the intervention)
dAA7[1]: *"Not all X are Y"* is equivalent to *"Some X are Y"*	Initial assumption (during the intervention)
dAA7[2]: "Not all X…" is different from "No X…"	Initial assumption (during the intervention)
dAA7[10]: The negation of *"All X are Y"* is *"Some X are not Y"*	New assumption (during the intervention)
dAA7[19]: The negation of *"All X are not Y"* is *"Some X are Y"*	New assumption (during the intervention)

Andrea's initial apparent link between the simple implicit negation of a UAS and its falsity

Andrea's initial implicit assumption about the relation between *"Not all X are Y"* and the falsity of *"All X are Y"* was in harmony with the mathematical perspective; that is, the statement *"Not all X are Y"* means that *"it is false, or it is not the case that all X are Y"* (her assumption dAA1[4]). Andrea used this assumption, for instance, to suggest that an argument was *repetitive*[130].

Specifically, during Discussion 1.1 Andrea explicitly pointed out that the argument *"because not all divisions of natural numbers are exact divisions"* to justify that statement St16 was false, was repetitive.

St16: *All divisions of natural numbers are exact divisions*

This shows that, at least implicitly, she equated the expressions *"it is false that all divisions of natural numbers are exact divisions"* with *"not all divisions of natural numbers are exact divisions"*.

Andrea's assumption was more explicit during Discussion 7.2 as she answered that the statements *"Not all divisions of natural numbers are exact divisions"* and *"It is not true that all divisions of natural numbers are exact divisions"* were not different from each other[131]. Nonetheless, her apparent awareness did not have any influence on the development of her understanding related to the negation of universal statements. This was revealed in her negation of universal statements, which was far away from its mathematical counterpart, as I show next.

Andrea's initial general approach to negate a UAS: Her "separate and substitute" approach (SS-approach)

Above I showed that Andrea used (at least implicitly) the link between negation and falsity of a UAS (her assumption dAA1[4]). Here I show that Andrea did not use this linkage to further reason about the negation of UASs. Instead, she used a different approach when finding equivalent statements for the implicit negation of a UAS.

In order to find an equivalent statement for the simple implicit negation of a UAS Andrea initially relied on her interpretation of "not all" as interchangeable with "some". In fact, Andrea's initial interpretation was compatible with the meaning of "some" in everyday language, namely, "some" means *some, but not all* (Epp, 2003; Grice, 1975), and I presume her assumption stemmed from this.

Discussion 7.1 formally introduced *Existential Statements*. In particular, we began the discussion by debating about the difference between statements St16 and St34,

St16: *All divisions of natural numbers are exact divisions*

St34: *Not all divisions of natural numbers are exact divisions*.

Andrea answered that the statements were different because the first statement was universal, whereas the second was *existential*[132]. She explained that "not all" was the same

[130] Recall that I call a repetitive to an argument of the form *"because not all X are Y"* to disprove *"All X are Y"* (for details, see Episode 5 in Section I.2.1.1 above).
[131] For details on the task the teachers solved in Discussion 7.2, see Appendix D7.
[132] During the intervention we mostly used the expression "particular statement" ("proposición particular" in Spanish) instead of "existential statement" since this is the most common way this is expressed in

as "some". When requested to provide a more precise answer, she revealed her use of a substitution in meaning.

> Andrea: *Because that "not all" is "some".*
>
> I: *Some what exactly? Complete the statement. I mean, if you say that this is the same as "some", some what exactly?*
>
> Andrea: *Some divisions of natural numbers are exact divisions.*

Andrea switched the "not all" part of statement St34 with "some" and kept the second condition of the statement unchanged, in its affirmative form ("are exact divisions").

In abstract terms, she used the assumption that *"Not all X are Y"* is equivalent to *"Some X are Y"*, which I call her assumption dAA7[1]. In Dawkins and Cook's (2017) and Dawkins' (2017) terms, Andrea's initial interpretation of the simple implicit negation of the UAS *"All X are Y"* entailed a *semantic substitution* of "not all" with "some" (see Figure 27).

Not all X are Y ≡ **[Not all]** X are Y ≡ **[Some]** X are Y ≡ Some X are Y

Figure 27. Andrea's "separate and substitute" approach to find an equivalent statement for the simple implicit negation of a UAS, "Not all X are Y".

This reveals that Andrea's claim that the statements St16 and St34 were different since the former was universal and the latter was existential was based on her use of a *semantic-substitution* approach. In Andrea's case I call this approach "separate and substitute" (or SS-approach). Andrea's SS-approach consisted in separating the statement so that she could substitute one of the parts ("not all") for its, according to her, equivalent meaning ("some"). Given that she identified St34 with the statement *"Some divisions of natural numbers are exact divisions"* (St24), she regarded St34 as an existential statement.

Andrea's approach to negating a UAS certainly differs from mathematical convention. In mathematics, St34 (*"Not all divisions of natural numbers are exact divisions"*) is equivalent to *"Some divisions of natural numbers are not exact divisions"* (St25), which is not the same statement as the one Andrea provided. Note that the mathematical equivalent for St34 has a second condition in its negative form ("are *not* exact divisions"), which differs from Andrea's statement St24 which has an affirmative form ("are exact divisions").

In abstract terms, the simple implicit negation *"Not all X are Y"* is mathematically equivalent to the statement *"Some X are not Y"* (see Figure 28) and not to *"Some X are Y"* as Andrea suggested.

Not all X are Y = **Not** [all X are Y] = Some X are not Y

Figure 28. A mathematical equivalence for "Not all X are Y".

Andrea was aware that "not" in St34 played the role of a negator, as she explicitly pointed it out later.

Spanish; however, here I decided to use "existential statement" to avoid possible confusions (e.g., between a "particular statement" and a "specific statement").

Andrea: *The "not" negates it, that means, it is not that it is "none", right? It is "not all", that means, it can be "some".*

Andrea's remark goes against a common assumption that individuals tend to make in relation to negations, which is that the statement *"Not all X are Y"* is equivalent to *"No X is Y"*[133]; that is, they assume that the negation of a UAS is its *contrary* instead of its *contradictory* statement (Horn, 2001). In the case of Andrea, she explicitly rejected this possibility before she was even asked about it[134]. I call her assumption dAA7[2].

This extract shows that Andrea did not consider that "not" negated the complete universal statement *"All X are Y"*, as it does in mathematics. Instead, based on her assumption dAA7[1], it seems that she assumed that the negator only affected the universal quantifier "all".

No apparent inconsistencies between Andrea's initial assumptions and a rule to negate USs

Later, in task (e) of Discussion 7.1, the teachers were explicitly asked whether the statement St34 (*"Not all divisions of natural numbers are exact divisions"*) states[135] that *some divisions of natural numbers are exact divisions*[136]. Andrea nodded in agreement, which was consistent with her initial assumption that both statements were equivalent (her assumption dAA7[1]). As a way to introduce a conflict to Andrea's initial SS-approach to negate UASs (see Figure 27), I suggested the following rule for the negation of universal statements:

Given Rule: *The negation of a universal statement results in an existential statement and the negator also negates the consequent.*

Based on this rule Andrea determined that the expected equivalent statement (for St34) was St25, *"Some divisions of natural numbers are not exact divisions"*.

Andrea did not show any disagreement with my given rule. Normally, Andrea was not afraid of expressing her ideas during our discussions, which sometimes included her disagreeing; however, here she did not show any opposition. I presume that she did not oppose the rule because she found no inconsistencies with her current set of assumptions. One of her initial assumptions related to existential statements (ESs) supported her harmony with the *given rule*. Andrea initially assumed that *"Some X are Y"* implied that *"Some X are not Y"*[137].

After stating St25 is equivalent to St34, based on my given rule, Andrea went on to say:

Andrea: *But also, even though the negator didn't affect the consequent, it [St24] would be true anyways.*

When Andrea says *"even though the negator didn't affect the consequent"*, she refers to St24 (*"Some divisions of natural numbers are exact divisions"*), whose consequent is not negated (it is "*are exact divisions*") and that *"it [St24] would be true anyways"* as, in her

[133] For references and details on this issue, see Chapter 2, Section I.8.

[134] A task that included an explicit discussion about this issue was included in Discussion 7.1 (see task d, in Discussion 7.1, Appendix D7); however, Andrea made her comment before the task was even presented to the teachers.

[135] In previous discussions, I had already explained that our use of "state" would be similar to "is equivalent to" or "express/convey the same information as".

[136] For details on the tasks that were part of Discussion 7.1, see Appendix D7.

[137] For details on Andrea's assumptions about "some-statements", see Section II.1 below.

view, both St24 and St25 are equivalent. Based on her initial assumption about some-statements that *"Some X are Y"* implies *"Some X are not Y"*, Andrea assumed that it did not really matter which "some-statement" was equivalent to St34. She could always infer one (affirmative/negative) "some statement" from the other (negative/affirmative) "some-statement".

All this suggests that Andrea did not reject the *given rule* to negate universal statements because she did not find any apparent contradiction with her set of initial related assumptions. The assimilation she made of the given rule prevented Andrea from understanding that the negation of St16 (St16: *All divisions of natural numbers are exact divisions*) was not the existential affirmative statement she gave (*"Some divisions of natural numbers are exact divisions"*). For Andrea, all her initial related assumptions and the *given rule* could coexist without any conflict.

A potential conflict: A true UAS and its false negation

Andrea's initial approach to the negation of USs began to stumble after Discussion 7.1. The teachers were requested to provide an example of a universal statement and determine its negation. Gessenia provided the (true) universal statement St86:

St86: *All natural numbers that end in digit zero are multiples of 5*,

for which she herself determined the expected negation St87:

St87: *Some natural numbers that end in digit zero are not multiples of 5*.

Andrea looked hesitant. She did not stop staring at the whiteboard where both statements (St86 and its negation St87) were written. Andrea asked us to justify that Gessenia's negation (St87) is false. This suggests that the statement and its truth value introduced some conflict.

Andrea's conflict might have stemmed from the mismatch between the truth values of St87 and the negation that Andrea would have initially expected from her *separate and substitute* approach:

St91: *Some natural numbers that end in digit zero are multiples of 5.*

Andrea could have noticed that St87 is false but her St91 is true. These two different truth values would conflict with her initial assumption about ESs, that *"Some X are Y"* implies *"Some X are not Y"*. As the compatibility of her SS-approach and my given rule depended on this assumption, once she began doubting her assumption about ESs, her confidence in her SS-approach (her assumption dAA7[1]) might have been shaken.

Our attention was then directed towards the disproof of the statement St87 (*Some natural numbers that end in digit zero are not multiples of 5*) and we lost track of our main discussion: the negation of St86. Hypothetically, if we had explored Andrea's conflict, she would have understood at this point that her SS-approach was inadequate to negate a UAS. But in fact, neither Andrea's possible conflict related to her initial assumption for ESs was resolved nor did we finish our discussion about the negation of St86. Thus, it is not clear whether after this discussion she reorganized any of her initial assumptions for the negation of UASs. However, the process of change might have begun with this discussion.

A functional accommodation of the given rule to negate USs

I planned to explicitly address Andrea's initial assumption related to ESs in Discussion 7.4.2[138] and it was resolved then[139]. Andrea came to understand that if *"Some X are Y"* was true, it did not necessarily follow that *"Some X are not Y"* was also true. She became aware that there may be cases where the latter was false. Consequently, Discussion 7.4.2 was an important breaking point, which certainly had an impact on Andrea's initial approach to negating universal statements. Andrea could not use her SS-approach and expect to produce the same existential statement we obtained with my given rule to negate USs. The two resulting "some-statements" would be different.

Andrea's "Distribute, Separate and Substitute" approach (her DSS-approach)

Andrea no longer used her SS-approach to negate a US after Discussion 7.4.2. The resolution of her conflict with ESs allowed her to change her initial approach to negating USs. Andrea performed what Steffe (1991, p. 183) calls the *functional accommodation of a rule*. Her accommodation of the *given rule* I gave was a new rule that allowed her to negate both universal affirmative and universal negative statements (UNSs). That rule I call Andrea's *"Distribute, Separate and Substitute"* rule (DSS-rule or DSS-approach). It consists of distributing the negator "not" to both the antecedent and consequent, grouping or separating the negator as affecting the quantifier "all" preparing it for substitution, and substituting "not all" with "some". In that way the affirmative/negative consequent was also changed into a negative/affirmative consequent (see Figure 29).

1) Negation of "All X are Y" = **not** all X are **not** Y = [not all X] are not Y
   = some X are not Y
2) Negation of "All X are not Y" = **not** all X are **not** not Y = [not all X] are Y
   = some X are Y

Figure 29. Andrea's DSS-approach to negate universal statements.

She used this approach in Discussion 9. The teachers discussed the implicit negation of another ES (*"There do not exist divisions of natural numbers that do not have a remainder equal to 0"*). Andrea gave the statement *"Not all divisions of natural numbers have a remainder equal to 0"* (S) to further explain her observations[140]. She said:

> Andrea: *when I have not all, that "not" changes "all" and changes the middle [she means the consequent] ... in the first statement [S] is, "not all", "some", and over there the middle changed to "do not give zero" ["do not have remainder zero"]*

Her substitution of "not all" with "some" in her DSS-approach suggests that Andrea might not have understood that an affirmative "some-statement" does not necessarily imply the negative "some-statement". However, in my view it meant that Andrea came up with a shortcut in order to negate USs. Instead of bringing her understanding of ESs into question, to me it is clear that she did not understand why negation worked the way it did from a mathematical perspective. Hence, she needed to come up with a rule that produced negations that were consistent with mathematics.

---

[138] For the specific content of Discussion 7.4.2, see Appendix D7.
[139] For details on the development of Andrea's assumptions related to inferences with "some-statements", see Section II.1 below.
[140] For details on this discussion, see Section II.3 below.

Next, I illustrate Andrea's uses of her DSS-approach and then I explain why I think that they are examples of functional accommodations of the *given rule*.

Andrea used her DSS-approach during and after Discussion 7.5. It was observed in her negation of one UAS and, notably, two universal *negative* statements, a type of USs we had not previously discussed. During Discussion 7.5 Andrea showed no difficulties identifying that *"Not all natural numbers divisible by 2 are divisible by 3"* (St103) is equivalent to *"Some natural numbers divisible by 2 are not divisible by 3"* (St106).

When discussing the truth value of St103 and its justification, Andrea decided to analyze its equivalent "some-statement" St106. Andrea provided the number 8 to prove that St106 (and therefore St103) was true. This already suggests that she acknowledged that the negation (St103) was a *negative* "some-statement" and not an affirmative "some-statement"; otherwise, the justification would have been different from the case she presented (Andrea would have given a number divisible by 2 that was divisible by 3, instead of one that was not divisible by 3).

After Discussion 7.5 Andrea did not hesitate when negating the universal *negative* statement St104:

St104: *All natural numbers divisible by 2 are not divisible by 3*

She obtained the expected statement *"Some natural numbers divisible by 2 are divisible by 3"*. She also negated the imaginary universal *negative* statement St109:

St109: *All odd numbers are not Innova numbers,*

and correctly obtained *"Some odd numbers are Innova numbers"*. In all three cases Andrea exhibited no hesitation when producing the negation.

I believe that Andrea's new DSS-approach to negate USs (her assumption dAA7[10]) was simply a *functional accommodation* of the *given rule* provided during the intervention. To give some context, during Discussion 7.1, after I introduced the *given rule*, I attempted to support the teachers' understanding of why the *given rule* worked the way it did. With that purpose I resorted to the *negation of the meaning* (Dubinsky, Elterman & Gong, 1988). It mainly consisted in making connections between the negation of a US and its falsity in order to base the understanding of the former on the latter[141]. Even though the negation-falsity relation was implicit in Andrea's assumption dAA1[4], her initial non-mathematical approach to negating USs revealed that she was not completely conscious of the negation-falsity linkage. For example, after Discussion 7.1 Andrea called our attention to the justification for the negation of St86 (St86: *All natural numbers that end in digit zero are multiples of 5*, see above). As a way to establish a link between the negation of St86 and its falsity, I reminded the teachers of a simpler case we had seen during Discussion 7.1. I drew the teachers' attention to the relation between the negation of St16, *"Some divisions of natural numbers are not exact divisions"*, and what it was involved in showing that St16 was false (showing at least one division of natural numbers that was not exact). I pointed to what we had on the whiteboard (see Figure 30) and specifically drew their attention to the disproof of St16 and the proof of its negation. Andrea's spontaneous question, *"is it the same justification?"*, revealed her lack of understanding of the relation between the negation of St16 ("statement 1" in Figure 30) and its falsity. I presume that, because she did not find any inconsistencies within her set

141 Recall that at this point of the discussions we had already explored the falsity of USs (for details, see Section I.2 above).

of initial related assumptions at that moment[142], Andrea did not find a real reason to make further connections between negation and falsity. When some inconsistencies surfaced during Discussion 7.4.2, she accommodated the *given rule* without making connections between the negation of a US and its falsity. In particular, after Discussion 7.5 Andrea managed to correctly provide the negation of the UNS St104 (see above), even though she still looked astonished with my explanation that linked the negation of the universal negative statement with its falsity. This shows that Andrea only accommodated the *given rule* in her set of current assumptions in order to use it. She did not understand the relation falsity-negation, neither did she aim at verifying consistency within her set of current related assumptions. Notably, her current understanding about allowed inferences with "some-statements" was not compatible with her DSS-approach to negate USs. Andrea still substituted "not all" with "some" mechanically, which supports that she did not understand the negation of USs.

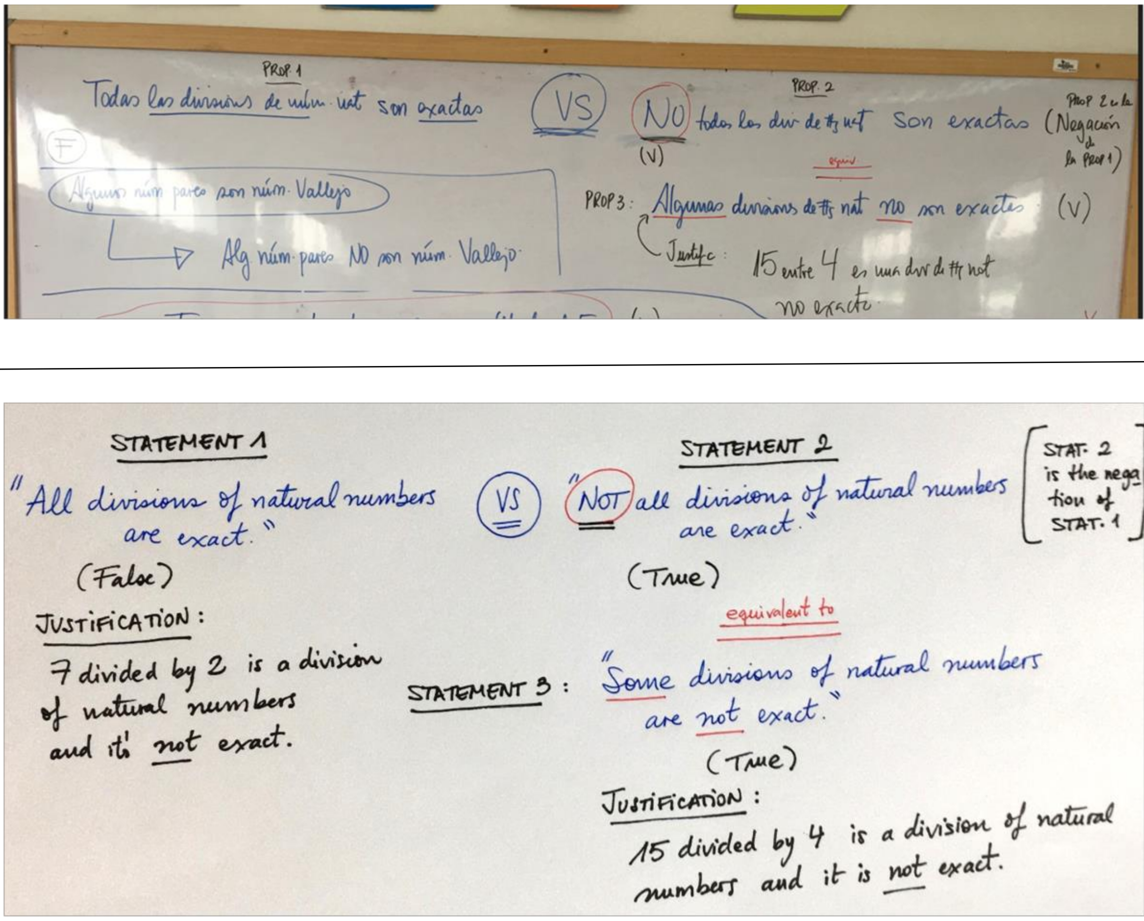


Figure 30. The notes on the whiteboard for Discussion 7.1 on the implicit negation of the UAS "All divisions of natural numbers are exact divisions" (Above the original version in Spanish; Below the translation into English, which includes a reconstruction of the justification for statement 1, which was spoken)

The case of Andrea shows that when integrating her assumptions about the negation of USs and her assumptions about "some-statements", the new set of assumptions was internally inconsistent. This means that her set of assumptions for "some-statements" changed and was consistent only locally (as I show in Section II.1 below).

[142] Recall that at that point of the intervention she still assumed that both (affirmative and negative) "some-statements" coexisted and so she used her SS-approach to negate USs.

Summary of Section I.4

Andrea began the intervention with two initial assumptions related to the negation of USs. The first was related to the linkage between negation and falsity: *"Not all X are Y"* and *"It is false that all X are Y"* state the same (her implicit assumption dAA1[4]). Andrea used her assumption to reject a *repetitive* argument and to point out that the simple implicit negation *"Not all X are Y"* was the same as *"It is not true that all X are Y"*. However, her initial assumption dAA1[4] did not play a significant role when she was expected to find an equivalent statement for *"Not all X are Y"* or explicitly negate *"All X are Y"*. Instead, Andrea used her initial assumption dAA7[1] to do that. Her assumption dAA7[1] was related to implicit negations of the form *"Not all X are Y"* and her semantic substitution of "not all" with "some". This led her to claim that *"Not all X are Y"* was equivalent to *"Some X are Y"* (her "separate and substitute" approach). This assumption is explained by Andrea's initial interpretation of "some" as "some, but not all", that stemmed from the everyday language context.

During the intervention I introduced a rule to negate USs (the *given rule*) which resulted in negations different from those Andrea would provide; however, it did not seem to disturb Andrea's initial SS-approach. While the *given rule* stated that *"Not all X are Y"* was equivalent only to *"Some X are not Y"*, Andrea's initial assumptions about "some-statements" still supported an apparent harmony with her SS-approach. At that moment Andrea still assumed that the statements *"Some X are Y"* and *"Some X are not Y"* implied each other; that is, it was not relevant to focus only on one of them when discussing the simple implicit negation *"Not all X are Y"*. This is the reason why Andrea did not reject the rule at that moment. When her initial assumption that *"Some X are Y"* implied that *"Some X are not Y"* was challenged, she quit using her SS-approach. Instead, she made a functional accommodation of the *given rule* and came up with a new rule, her "distribute, separate and substitute" (DSS-) approach. Her DSS-approach still involved her use of the substitution of "not all" with "some". This suggested that Andrea did not come to understand why the *given rule* to negate USs worked the way it did in mathematics.

Figure 31 shows the development of Andrea's assumptions about the negation of UASs and the elements or aspects of the intervention that might have supported or triggered changes[143].

[143] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

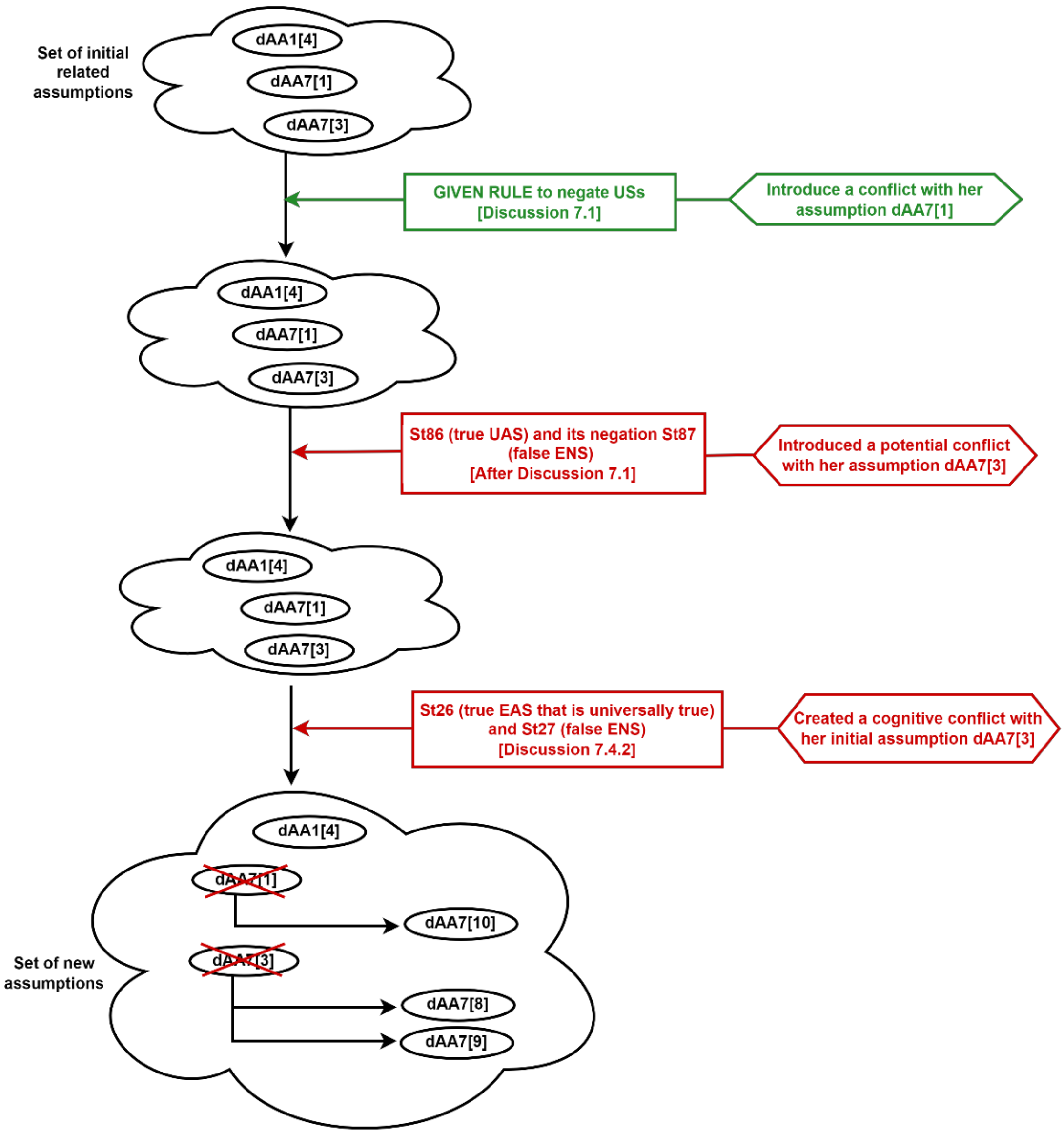


Figure 31. The development of Andrea's assumptions about the negation of UASs (dAA1[4]: "Not all X are Y" and "It is false that all X are Y" state the same; dAA7[1]: "Not all X are Y" is equivalent to "Some X are Y"; dAA7[3]: "Some X are Y" necessarily implies that "Some X are not Y"; dAA7[10]: The negation of "All X are Y" is "Some X are not Y" ("Some X are Y"); dAA7[8]: If "Some X are Y" and "All X are Y" are both true, then "Some X are not Y" is false; dAA7[9]: If "Some X are Y" is true and "All X are Y" is false, then "Some X are not Y" is true)

II. The teachers' assumptions about Existential Statements

This section is devoted to the development of Andrea' assumptions about existential – both affirmative and negative – statements (ESs)[144]. My exclusive focus on Andrea's assumptions is because of her high participation and explicit display of her thinking process. Gessenia was not consistent with her thinking nor she did not express clear reasons for why she held certain assumptions. Lizbeth was not present for Discussion 7, which is when existential statements were formally discussed. Hence it was difficult to have a complete overview of her thinking.

I focus on Andrea's assumptions related to three topics: (1) Establishing truth and proving existential statements; (2) Establishing falsity and disproving existential statements; (3) Negation of existential statements.

Before I dive into these themes, I list the discussions of the 2018-intervention that included any kind of debate related to existential statements as well as the aspects about them that might be relevant to this chapter. Existential statements were planned to be introduced during Discussion 7 as the negation of universal statements. Nevertheless, the existential quantifier *"some"* was already included much earlier, in Discussion 1.5, where the meaning *"at least one"* and *"one or more"* for *"some"* was prompted to be used in a task. Also, during Discussion 2.1 Andrea asked to discuss the truth value of an existential statement.

The existential statements seen during the 2018-intervention were, in general, of the form: *"Some X are (not) Y"*, *"There exists X that is (not) Y"*, and *"There is X that is (not) Y"*.

Now I proceed to expand on each of the assumptions that Andrea developed about the three aforementioned topics.

1. Establishing Truth and Proving Existential Statements

In this section I focus on the criteria Andrea used to establish truth of existential statements and the status of confirming examples when proving existential statements.

From a mathematical perspective, one confirming example is sufficient justification to prove an existential statement[145]. Andrea's starting assumption was not aligned with that perspective. Additionally, her initial interpretation of the quantifier "some" supported immediate inferences with "some-statements" that were mathematically invalid. Likewise, her initial interpretation of "there is" made that Andrea distinguished between "some-statements" and "there-is-statements". Here I show the role that the intervention played in Andrea's reconsideration of some of her initial assumptions.

Table 16 includes Andrea's assumptions related to the truth of ESs and their proving, as they were first observed during the intervention.

[144] For details on Existential Statements, see the Mathematical Framework in Chapter 3, Section II.

[145] For details on the mathematical aspect, see Chapter 3, Section II.3.

Table 16. Andrea's assumptions about establishing truth and proving ESs

Assumption	**Type of assumption (when first observed)**
bAA6: *"There is X that is Y"* does not convey quantity	Initial assumption (before the intervention)
bAA7: The true US *"All X are Y"* implies that *"There is X that is Y"* is true	Initial assumption (before the intervention)
dAA1[1]: The true US *"All X are Y"* implies that *"Some X are Y"* is false	Initial assumption (during the intervention)
dAA1[8]: *"Some"* means *"from all, one group, but not all"*	Initial assumption (during the intervention)
dAA7[3]: *"Some X are Y"* necessarily implies that *"Some X are not Y"*	Initial assumption (during the intervention)
dAA1[10]: Confirming examples are insufficient to prove a statement (as a justification, ONLY for false universal statements)	New assumption (during the intervention)
dAA7[4]: One confirming example is sufficient to prove a "some-statement"	New assumption (during the intervention)
dAA7[7]: "Some" means "at least one, and maybe all"	New assumption (during the intervention)
dAA7[8]: If *"Some X are Y"* and *"All X are Y"* are both true, then *"Some X are not Y"* is false	New assumption (during the intervention)
dAA7[9]: If *"Some X are Y"* is true and *"All X are Y"* is false, then *"Some X are not Y"* is true	New assumption (during the intervention)

I begin by focusing on Andrea's initial assumptions related to establishing and/or proving the truth of ESs. I also consider one of my inputs, which is the mathematical meaning of "some" I prompted the teachers to use during our discussions. I continue with two conflicting situations that seem to have supported Andrea's new insights about existential statements in this context. I finish with some evidence of Andrea's emergent understanding, which takes a clearer form through her teaching as she draws her students' attention to some important factors that might explain her own understanding.

Andrea's initial related assumptions

Andrea displayed five initial assumptions about existential statements. Two of them (assumptions bAA6 and bAA7) were about "there-is-statements" and emerged during the First Exploratory Interview, before the intervention; whereas three (assumptions dAA1[1], dAA1[8] and dAA7[3]) concerned "some-statements" and emerged during the intervention. Another assumption (dAA1[10]) was an overgeneralization that Andrea made during the first part of the intervention. It is considered here because it is one assumption that involved the proving of ESs and emerged before the proof-related discussions took place[146].

The statement "There is X that is Y" does not convey quantity (bAA6)

Before the intervention for teachers Andrea used her initial assumption that existential statements of the form *"There is X that is Y"* did not convey any specific quantity. This was manifested through her analysis of a "there-is-statement" during the First Exploratory Interview. The teachers were asked to solve a task that consisted of the UAS St2,

[146] Andrea's assumption dAA1[10] was also discussed in Section I.3.1.1 above.

St2: *All «Vallejo» numbers are even numbers.*

Based on St2 the teachers were requested to choose from a list of twelve statements those that conveyed the same information as St2. Andrea chose the four following expected statements:

- *"If a number is «Vallejo» then it is an even number"* (option c),
- *"Every «Vallejo» number is an even number"* (option g),
- *"The «Vallejo» numbers are even numbers"* (option h), and
- *"There are no «Vallejo» numbers that are not even"* (option k).

However, Andrea showed hesitation about options "e" (*"The even numbers are not «Vallejo» numbers"*)[147] and "j" (*"There are «Vallejo» numbers that are even numbers"*) and whether they should be chosen or not. As Andrea explained why she hesitated about statement "j", she revealed two of her initial assumptions about "there-is-statements".

> Andrea: *I'm not sure about "j", because it says, there are Vallejo numbers that are even numbers. All [Vallejo numbers] are even. But it also holds that there are Vallejo numbers that are even numbers. All [Vallejo numbers] are even, but if I choose any of them, I mean, it is even, too. In other words, here [statement "j"] there is no statement that conveys quantity. That is why this [statement "j"] could be too.*

Andrea explicitly stated that the existential statement "j" did not provide quantity (*"here [statement "j"] there is no statement that conveys quantity"*). It supposes that she found that the quantifier "there is" did not specify a concrete number of cases involved in the statement (*one of them, two of them, all of them* or, otherwise, *how many of them*).

Her initial assumption about the quantity of a "there-is-statement" might be explained in terms of the quantifier "there is" and the way it is used in everyday-language conversations. The meaning of the quantifier "hay" in common-language Spanish ("there is" in English) is the same as "there exists". "There is" does not have a connotation that conveys precision. It conveys *existence*, though not a real specification on the number of objects it refers to. This might have influenced Andrea to hesitate on choosing statement "j" given that it might have encompassed all "Vallejo" numbers.

The true statement "All X are Y" implies that "There is X that is Y" is true (bAA7)

Andrea presumed that the existential statement "j" was also a possible answer because she seems to have been focused on it as an immediate inference from the given universal statement St2. Based on her assumption that St2 was true[148] and her apparent implicit use of a principle of logic known as *universal instantiation*, she inferred that the existential statement "j" was true (her assumption bAA7).

> *Universal Instantiation:* If a property is true of *everything* in a set, then it is true for any particular element of the set. (Epp, 2020, p. 146)

Her use of this principle is manifested by her assertion that *"All [Vallejo numbers] are even, but if I choose any of them, I mean, it is even too"*, which suggested her immediate inference that the existential statement "j" was true. Andrea seems to have hesitated about

[147] The discussion about option "e" was already included in Section I.1.1.1.
[148] In Section I.1.1 I explained that the teachers had the tendency to assume that given an imaginary statement, it was a true statement.

choosing or not option "j" because the task was about equivalent statements and not inferences.

"Some" means "from all, one group, but not all" (dAA1[8])

Even though it was not the main goal of Discussion 1 of the intervention for teachers to engage in a thorough discussion about existential statements, the term *"some"* was included in one of the tasks of Discussion 1.5[149]. I asked the teachers what they understood by "some". Andrea explained that she understood it as *"from all, one group"*.

In Discussion 7.1 Andrea complemented her initial interpretation of "some" with her remark that "some" and "not all" were the same[150]. Her observation complemented her initial meaning of "some" as *"from all, one group, but not all"* (her assumption dAA1[8]), which is compatible with the interpretation for "some" in everyday conversations (Grice, 1975; Epp, 2003).

As in the case of the existential quantifier "there is", Andrea's initial interpretation for "some" does not convey a specific quantity. Her expression "one group" does not specify the exact number of members in the "group" she refers to. However, unlike the case of "there is", "some" discarded the possibility of "all" ("some" means "not all").

"Some X are Y" necessarily implies that "Some X are not Y" (dAA7[3])

Andrea explicitly used her initial assumption that *"Some X are Y"* necessarily implies that *"Some X are not Y"* during Discussion 7.1 (her assumption dAA7[3]). As I explained before (in Section I.4 above), based on a rule I gave to negate universal statements I expected that the teachers would realize that the statement *"Not all divisions of natural numbers are exact divisions"* (St34) was equivalent to St25:

> St25: *Some divisions of natural numbers are not exact*,

and not to St24:

> St24: *Some divisions of natural numbers are exact.*

Andrea's reaction consisted in suggesting an inference that seemed obvious to her (dAA7[3]).

> Andrea: *But also, even though the negator didn't affect the consequent, it [St24] would be true anyways.*

With her claim that *"the negator didn't affect the consequent"* Andrea meant that the consequent in the negation she had initially suggested (St24) was not in a negative form, in contrast to the negation St25 produced with the *given rule*. With her complete utterance she tried to explain that it did not matter that the negation was St25 instead of St24, because St24 was true in any case. That is, Andrea brought up her initial assumption dAA7[3] that both existential statements (here St24 and St25) necessarily coexisted, based on the initial meaning she used for the quantifier "some". Andrea downplayed the negation with the *given rule* by taking for granted that the other existential statement *"would be true anyways"*.

---

149 For details on the tasks included in Discussion 1.5, see Appendix D1.
150 For details, see Section I.4 above.

In other words, Andrea initially assumed that the true existential statement *“Some X are Y”* implied that one part of *X* satisfied the condition *Y*, and the other part of *X* did not, as I represented in Figure 32.

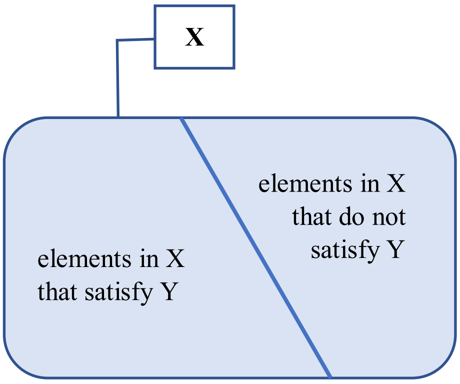


Figure 32. Andrea's initial interpretation of the true ES "Some X are Y"

As a result of her initial interpretation for “some” Andrea assumed that *“Some X are Y”* necessarily implied that *“Some X are not Y”* (her initial assumption dAA7[3]).

A true US “All X are Y” implies that “Some X are Y” is false (dAA1[1])

Determining the truth value of a “some-statement” was directly associated with Andrea’s initial interpretation of the quantifier “some”. Particularly, she assumed that a “some-statement” was false if the statement was universally true (her assumption dAA1[1]). Evidence of Andrea’s assumption arose in an activity the teachers solved right before Discussion 1 of the intervention for teachers. The teachers were given 18 statements and were asked to group the statements according to the criteria they decided to choose. The teachers opted to separate the statements by their truth value. Among the false statements they included the (true) existential statement St26:

St26: *Some numbers divisible by 4 are even numbers.*

The teachers were somehow aware[151] that St26 was universally true; that is, the existential statement was actually true for all the cases involved in the statement (the numbers divisible by 4). The fact that the teachers selected St26 as one false statement was the first indication that the teachers, and Andrea in particular, used the assumption that both statements of the form *“Some X are Y”* and *“All X are Y”* could not be true.

An explanation for Andrea’s initial assumption dAA1[1] could be given in terms of her initial interpretation of the quantifier “some”, which discarded the possibility of “all” (her initial assumption dAA1[8]). Andrea assumed that since *“All numbers divisible by 4 are even numbers”* was true, then *“Some numbers divisible by 4 are even numbers”* should be false.

Andrea’s initial assumption dAA1[1] might seem contradictory to her initial assumption bAA7. Whereas Andrea assumed that from a true US *“All X are Y”*, *“Some X are Y”* was false, she admitted that from the same true US, the statement *“There is X that is Y”* was true. This, from a mathematical/logical view does not make any sense given that *“Some X are Y”* and *“There is X that is Y”* are equivalent statements. However, I explain Andrea’s apparent inconsistency with the distinction she made between “some-

[151] We did not engage in a discussion for the justifications of the truth values they assigned to the statements in this activity.

statements" and "there-is-statements". For Andrea the statements meant different things. On one hand, as I explained above, Andrea linked "there-is-statements" with *existence*; that is, *"There is X that is Y"* suggested the existence of element(s) in *X* that satisfied *Y*, which was not in conflict with *"All X are Y"* since Andrea was aware that based on the true US, the statement held for *every* existing element in *X*. On the other hand, her initial meaning for "some" as "from all, one group, but not all" prevented her from acknowledging the possibility of having a "some-statement" that was universally true. For Andrea, the fact that a statement of the form *"Some X are Y"* was true implied that the respective negative statement *"Some X are not Y"* was also true. Whereas a "there-is-statement" did not deny the possibility of being universally true, for Andrea a "some-statement" definitely did. Andrea's distinction of the two forms of existential statements ("there is-statements" and "some-statements") became clearer as we discussed the negation of ESs (for details, see Section II.3 below).

Agreeing on the meaning of "some" as "at least one"

Instead of rejecting Andrea's initial interpretation for "some" (assumption dAA1[8]) during Discussion 1.5, I asked the teachers to use the mathematical meaning of "some" in our discussions; namely, "some" means "at least one". Andrea did not object to my request to adopt this meaning, presumably because at that moment she found no apparent contradiction with her initial meaning for it.

Andrea's reconsideration of her assumption dAA1[10]: A Conflict

Andrea's initial assumption dAA1[10] related to ESs was her first assumption that changed. During the intervention Andrea faced a conflict when one of her assumptions (*confirming examples were insufficient to prove*, her assumption dAA1[10], see Section I.3.1.1) was challenged. The conflict arose as Andrea learned that an example may indeed be sufficient to prove an existential statement.

The Conflicting Example: Proving true "some-statements"

One of the tasks in Discussion 2.1 included the "some-statement" St46:

St46: *Some FWM distributions have a remainder zero.*

Even though the task did not include a discussion about the truth value for St46, Andrea drew our attention to it as she asserted that St46 was true[152]. I asked the other teachers whether they agreed with Andrea's claim. Gessenia agreed and provided the confirming example *10 divided by 2* to explain why the statement was true. Andrea looked puzzled. I raised the sufficiency of Gessenia's justification to prove St46. With the goal of supporting the discussion, I reminded the teachers the mathematical meaning of the quantifier "some" ("at least one", which I had introduced as an input in Discussion 1.5) and on this basis we analyzed what it meant for St46 to be true. Andrea's conflict stemmed from her unexpected mismatch between her assumption dAA1[10] that *examples were insufficient to prove* and her emergent realization that one example was indeed sufficient to prove the "some-statement" St46. A clear manifestation of the conflict Andrea

[152] For details on the content of the task, see Appendix D2.

experienced at that moment was not only expressed by her gestures suggesting surprise, but also by the question she asked afterwards.

> Andrea: *So, over here, is at least one [example] sufficient?*

Her question and surprise exhibited her uncertainty about the kind of mathematical evidence that proved St46 (*"is at least one [example] sufficient?"*). Moreover, based on the expression she used (*"So, over here…"*), Andrea made an explicit contrast with what she seems to have assumed before this discussion.

Presumably, the main reason why Andrea claimed that St46 was true in the first place was based on her awareness that *one group of FWM distributions had a remainder zero* and *the other group of FWM distributions did not have a remainder zero*, which involved the use of her initial meaning for "some". That means that she seems to have initially determined that St46 was true by relying on a *wrong frame* (Meissner, 1986[153]), which was her daily-life frame.

As part of the follow-up discussion, the quantifier "some" and its mathematical meaning were highlighted again to explain why a confirming example was sufficient to prove St46. Lizbeth contrasted the justification for this case and the justification for the false universal statements we had seen in Discussion 1. The discussion finished and Andrea did not participate again, so I cannot say whether her conflict was resolved or not.

Much later during Discussion 7.1 instead of evaluating whether a (true) ES was true Andrea opted to focus on its falsity, even though she was uncertain of it. The discussion was centered on the statements St16 and St34 and whether they made the same claim.

> St16: *All divisions of natural numbers are exact divisions*
>
> St34: *Not all divisions of natural numbers are exact divisions*

Both Andrea and Gessenia[154] agreed that St16 and St34 did not state the same thing. In order to further explain their answers Andrea argued that the statements' *quantity* (Copi, et al., 2014) was different (*"this [pointed to St16] is universal, while the other [St34] is existential"*); whereas Gessenia referred to the different truth value to support her answer (St16 is false and St34 is true), to which Andrea agreed.

To explain that St34 was an existential statement, Andrea relied on her initial meaning for *"some"*: *"not all" is "some"* (her assumption dAA1[8]). Andrea assumed that St34 meant the same thing as statement St24[155]:

> St24: *Some divisions of natural numbers are exact divisions.*

At that time, she held assumption dAA7[5a]: *"Some X are Y" is false if it has a "counterexample"* (see Section II.2 for details). As she could think of divisions that were not exact, she concluded that St24 is false, which led her to reconsider the truth value of St34.

> Andrea: *I believe this [St34] would be false as well. Well, the thing is that we haven't seen, um, this sort of cases, because I-, I can pick a counterexample there… this [St24] would be false because I can pick one [example] that doesn't hold.*

---

[153] For details on Meissner's (1986) concept "wrong frame", see Chapter 2, Section III.2.
[154] Recall that Lizbeth did not attend Discussion 7 of the intervention.
[155] For details on this issue, see Section I.4 above.

Andrea's claim that statements St34 and St24 were the same was stronger than her initial claim that St34 was true. As she now claimed that St24 was false (*"this [St24] would be false"*) she changed the truth value of St34. Andrea concluded that St24 was false since she could think of a counterexample for it (*"because I can pick one [example] that doesn't hold"*)[156], although Andrea was not completely sure (*"Well, the thing is that we haven't seen that, this sort of cases"*).

If Andrea had considered that the "some-statement" St24 could be true (what was entailed to prove it mathematically), she would have probably concluded that the statement was true, as Gessenia did next. Gessenia's explanation (*"because at least one division of natural numbers is exact"*) was based on the mathematical meaning for the quantifier "some" and the evidence she presented (20 divided by 4) was aligned with it. Andrea did not show any refusal of this. Instead, Andrea's immediate response suggested that she accepted Gessenia's conclusion. Andrea changed again the truth value of St34 (from false to true again), as she still assumed that St34 and St24 were equivalent and now she had accepted that St24 was true.

This means that depending on where Andrea's attention was placed, Andrea modified her decision about the truth value for St34. In this case Andrea's attention was focused on St24 and her initial assumption that St34 and St24 were equivalent. Andrea shifted the truth value of St34 because she shifted first the truth value of St24. She took a doubtful path to determine the truth value of St24. For some reason Andrea overlooked the possibility of having a true "some-statement" (St24), which could have avoided her struggle to determine why St24 was false as her main approach. Once the teachers' attention was directed to the logical interpretation of St24 and what was required to prove it, Andrea finally concluded that St24 was true.

After this discussion Andrea had several opportunities to prove (both affirmative and negative) "some-statements", which she did without hesitating, by providing one confirming example. This suggests that Andrea had already adopted the mathematical meaning of "some" to prove "some-statements", instead of using her initial meaning (dAA1[8]).

Evidence of Andrea's understanding of proving "some-statements": Her Teaching

After the intervention for teachers, Andrea focused on three main aspects when her teaching included proving "some-statements": (a) the meaning of the quantifier *"some"*; (b) its contrast with the quantifier *"all"*; (c) the sufficient justification to prove a "some-statement".

The first opportunity Andrea had to introduce the existential quantifier *"some"* to her class was during her teaching of Session 7. The whole class was engaged in solving a task that consisted in finding out whether the (false) "some-statement" St131 was true and explaining why.

St131: *Some divisions by 4 have a remainder equal to 7*

This was the first existential statement that Andrea's class discussed, which basically meant that at least at a first stage each student used his/her own initial assumptions about them. The way Andrea supported her students' understanding revealed Andrea's current

[156] For details on this issue, see Section II.2 below.

personal understanding. Andrea explained the existential quantifier *"some"* as meaning *"at least one"* and focused on contrasting it with the universal quantifier *"all"*.

> Andrea: *Is there indeed a division like that? When I say "some", I mean "at least one". When I divide by 4, is the remainder going to be equal to 7? Let's see, why?... Here, were you asked for ALL divisions?... It means, if there was at least one, if one holds, Giussepe [calling for one student's attention], if one holds, this is valid. At least one division, it says, one division by 4, that has a remainder equal to 7. Let's see whether this is true or not… there will be some divisions, here it says, such that when you divide by 4, the remainder will be equal to 7? There will be some?*

Andrea's question *"Here, were you asked for ALL divisions?"* clearly showed that she wanted her students to notice the difference between *"some"* and *"all"*. Andrea also highlighted the kind of evidence that was expected to be shown in order to accept that St131 was true (*"if there was at least one, ..., if one holds, this is valid"*). To do so, she drew attention to the quantifier *"some"* and its meaning (*"When I say 'some', I mean 'at least one'"*). In addition, she specified the characteristics of the division that was expected in order to (hypothetically) prove the statement (*"At least one division, it says, one division by 4, that has a remainder equal to 7"*). Andrea challenged her students to find at least one division with these characteristics in order to prove that the St131 was true[157].

The same kind of teaching pattern was used by Andrea during her teaching of Session 10. The class was expected to determine the truth value of the (true) "some-statement" St142 and justify it.

> St142: *Some natural numbers are divisible by 4*

Andrea began the feedback by drawing her class's attention to one student's incorrect answer. The student (Victor) had claimed that the statement was false and used the example *1 divided by 4* to justify his answer.

> Andrea: *Let's see, he tells me, miss, 1 is not divisible by 4 and it is a natural number. Let's see, 1 divided by 4, what is it equal to?... ah, miss, 1 is a natural number and it does not satisfy that it is divisible by 4. Now, I ask you Victor, here, does it state that ALL natural numbers are divisible by 4?... and what does "some" mean?... "some" means AT LEAST ONE. It does not mean at least two. If there is one example, it does hold. If there is at least one example, it does hold, this is true… 4 is divisible by 4, sufficient!... you-, over there they say, miss, 8 is divisible by 4, sufficient! If you found at least one example, then it does hold. This is true.*

This episode is interesting for at least two reasons. First, St142 was a true "some-statement"; however, one student claimed that the statement was false. He provided what Andrea herself called a "counterexample" during the intervention for teachers (see Andrea's explanations for the truth value of statements St34 and St24 during Discussion 7.1 above). This time Andrea was aware that such an example did not disprove the statement. To support her student's understanding she asked the student whether the statement referred to "all" instead since she knew the student's example was actually a counterexample for the respective "all-statement". Andrea then drew his attention to the meaning of "some" so that the student reconsidered the statement to be true, which she did not consider herself when she analyzed the truth value of the "some-statement" St24 during Discussion 7.1 (see above). Second, Andrea explicitly emphasized the sufficiency

[157] For details on Andrea's class disproving of St131, see Section II.2 below.

of showing one confirming example to prove that St142 was true. To do that, Andrea used expressions such as *"If there is at least one example, it does hold, this is true"*, and *"If you found at least one example, then it does hold. This is true."*. Moreover, she was very explicit about the sufficiency of at least one supporting example when she said *"4 is divisible by 4, sufficient!... you-, over there they say, miss, 8 is divisible by 4, sufficient!"*.

Andrea's reconsideration of her initial assumption dAA7[3]: A Cognitive Conflict

Andrea's initial assumption dAA7[3], that *"Some X are Y"* necessarily implies that *"Some X are not Y"*, was the second assumption she questioned. Andrea's initial assumption dAA7[3] was still exhibited during discussion 7.4.1, where the focus was on the two "some-statements" St24 and St25,

> St24: *Some divisions are exact,*
>
> St25: *Some divisions are not exact.*

These were essentially the same "some-statements" discussed in Discussion 7.1. So, Andrea identified that St24 was true and provided a confirming example to prove it. Andrea properly observed that St24 does not state St25. Nevertheless, her subsequent comment, though with traces of uncertainty, suggested that her initial assumption dAA7[3] was resistant to change.

> Andrea: *But this is going to always hold, isn't it?*

With *"this"* Andrea meant the negative "some-statement" St25. Specifically, she attempted to say that given a true affirmative "some-statement", the respective negative "some-statement" always held to be true as a consequence (her assumption dAA7[3]).

I gave back the problem to the teachers. I asked the teachers whether the fact that St24 was true implied that St25 was also true. Andrea's answer again exhibited traces of her initial assumption dAA7[3].

> Andrea: *Mathematically not, but logically, yes.*

With her claim, Andrea revealed a distinction she made about what we discussed from a mathematical point of view and what she still expected from a "logical" perspective. From a mathematical point of view, she knew that St24 did not imply St25; however, from a "logical" perspective, she still believed that it did. Andrea's use of the term *"logical"* was most likely linked to what a person might have expected from common sense (daily-life reasoning). The everyday-language interpretation for "some" (dAA1[8]) supported her "logical" assumption dAA7[3].

The teachers were asked to reflect on whether St25 was true given that St24 was true.

> Andrea: *No, they are independent… but I can deduce one [statement] from the other one.*

Andrea's use of the verb *"deduce"* was most likely associated with her personal certainty about her assumption dAA7[3]. Hence, Andrea's use of the terms *"logical"* and *"deduce"* were not of mathematical nature.

Her claims clearly showed that even at this point of the intervention Andrea based her reasoning related to "some-statements" on a *wrong frame* (Meissner, 1986,), daily-life context.

The Conflicting Example: A true "some-statement" that is universally true

In Discussion 7.4.2 the teachers were given the following two "some-statements":

St26: *Some numbers divisible by 4 are even numbers*

St27: *Some numbers divisible by 4 are not even numbers.*

Before we began to answer a sequence of questions about these two statements, I made sure that all the teachers were aware that the statement St26 was universally true; that is, the teacher became aware that *"All numbers divisible by 4 are even numbers"* (St92) was true, although we only shared hints of a proof for it.

St92: *All numbers divisible by 4 are even numbers*

The first task consisted in determining whether St26 was true or not and explain why. Andrea correctly claimed that St26 was true and provided the number 12 to support her answer. The second task requested the teachers to answer whether statement St26 stated the same as St27. All the teachers agreed that it did not.

Given that in the first task the teachers determined that St26 was true, the third task asked the teachers whether the fact that St26 was true implied that St27 was also true. Notice that unlike in previous examples, here the negative "some-statement" (St27) was false, which was planned with the goal of breaking the pattern of reasoning that at least Andrea was used to. Andrea shook her head in disagreement. Attention was then drawn to the truth value of St27, which Andrea quickly identified as false.

Andrea's responses to the Cognitive Conflict

Andrea showed four different responses to the conflict she experienced in Discussion 7.4.2, which revealed her insights. Her *first response* consisted of her formulation of a generically-formulated claim, which was based on her observations from the conflicting example.

> Andrea: *When the existential statement is also a universal statement, because that existential statement [St26] is also universal, that "some" is universal, okay, then it [St26] has nothing to do with its "little sibling" [she meant St27]. But if it [St26] is not universal, then it does, because when it is not universal, it means that there is a group that does satisfy [the second condition] and another group that does not [satisfy such condition].*

This is interesting as Andrea noticed that when the affirmative "some-statement" was universally true (or as she put it, *"when the existential statement is also a universal statement"*), then the respective negative "some-statement" was not true (or as she put it, *"then it [St26] has nothing to do with its "little sibling" [she meant St27]"*). On the other hand, when the affirmative "some-statement" was true, but not universally true (or as Andrea put it, *"if it [St26] is not universal"*), the respective negative "some-statement" was true (or in Andrea's terms, *"it means that there is a group that does satisfy [the second condition] and another group that does not [satisfy such condition]"*).

Even though Andrea based her reasoning on the examples we analyzed in Discussion 7.4.2, the way she formulated her observations attempted to move beyond to make a general observation. Andrea's remarks became part of her new assumptions about "some-statements" and can be summarized in general terms as follows:

i) If *"Some X are Y"* and *"All X are Y"* are both true, then *"Some X are not Y"* is false.

ii) If *"Some X are Y"* is true and *"All X are Y"* is false, then *"Some X are not Y"* is true.

Her new assumptions explicitly and precisely stated the conditions for when "some-statements" were true or false in relation to the truth value of "all-statements". In contrast to her initial assumption dAA7[3], her new assumption (i) admitted cases where *"Some X are not Y"* may be false even though *"Some X are Y"* is true (her new assumption dAA7[8]). Andrea's new assumption (i) implicitly suggested a refinement in Andrea's initial assumption dAA1[8] (*"some"* means *"from all, one group, but not all"*). It allowed that "some" might mean "all" (her new assumption dAA7[7]). In comparison with her initial assumption dAA1[1], her new assumption (i) also opened the possibility for statements *"All X are Y"* and *"Some X are Y"* to be both true.

Andrea's *second response* to the conflict was her production of a familiar-context example that satisfied the conditions of the conflicting example (below I go into more detail about her example). Andrea came up with the following example spontaneously.

> Andrea: *For example, I can say, here in this classroom I know everybody got an "A" in a test, but it is also valid to say that some got an "A", and it does not mean that some did not get an "A".*

Her *third response* comprised her explicit comparison of what she just learned from the conflicting case with the way she knew that the quantifier *"some"* was typically used and taught.

> Andrea: *But, when we work with quantifiers, we don't use "some" if there's already an "all" [she means, if it holds for all].*

Notably, speaking about quantifiers and connectives, Lee and Smith (2009) explained that, "from a linguistic perspective, the uses of these terms often follow the maxims of pragmatics in everyday conversation" (p. 22). In this context, Lee and Smith pointed to the *quantity* maxim of conversation (Grice, 1975) to explain the use of "some" in everyday conversations: "It is also informative to the best of the speaker's knowledge, or 'all' would be used instead" (p. 22). Andrea's comparison precisely referred to this idea. For her it was strange to use "some" when the statement held for "all" given that her initial interpretation of "some" was clearly framed in the common-life context; that is, it was non-mathematical.

Andrea still exhibited a *fourth response* to the conflict she faced. Her sort of "out loud reflection" consisted in using our own context (at that moment we were only four female teachers in the classroom) to shape a new example for which, as she put it, she would have never used the quantifier *"some"* before.

> Andrea: *But, is it valid to say that-, imagine that somebody comes in and asks us, here, are there teachers who are women? ... even though we are all women?... I thought "some" was only used ... when "not all" holds.*

Andrea's responses to the conflict did not only exhibit her evident realization of the conflict, but they also involved a modification of Andrea's set of initial assumptions related to ESs and, particularly, to "some-statements".

In summary, Andrea indirectly announced her new assumptions through a summary of her observations (her first response). She built coherence between the examples she produced and her new set of assumptions (her second and fourth responses). Andrea also contrasted the conflicting assumptions (her third response). Andrea's responses were similar to Chan, Burtis and Bereiter's (1997) *explicit knowledge building* and Chinn and Brewer's (1993) *change of theory* as they revealed major changes in her set of assumptions related to "some-statements"[158]. Andrea's initial assumption dAA7[3] changed and this change triggered the change of her related assumptions dAA1[1], dAA7[3] and dAA1[8].

Some remarks about the characteristics of the examples produced by Andrea

As a result of the cognitive conflict that Andrea experienced during Discussion 7.4.2, she produced two examples of "some-statements" that revealed her own current understanding of inferences that may or may not be possible with existential statements.

Example 1: *Some students got an "A" in a test* (when all got an "A")

Example 2: *Some teachers in the classroom are women* (when all teachers were indeed women)

These examples have some characteristics that are worth being pointed out. One is associated with the conditions satisfied by Andrea's examples. The two examples are true affirmative "some-statements" that are universally true and therefore the respective negative "some-statements" are false, which are essentially the same characteristics of the conflicting example. Likewise, the two examples were constructed during the intervention for teachers and, notably, right after Andrea formulated a generic claim for her observations about the conflicting example. Much later, after the intervention, Andrea recalled and used her second example, which revealed a "long-term" availability of her insights about "some-statements". Moreover, that exhibited the paradigmatic nature of Andrea's examples.

Another feature is linked to the context. While Andrea's examples belong to a real-life context, the conflicting example was framed in a mathematical context. The first example was about scores in a test, while the second was about people in a classroom. About the context aspect, some researchers (e.g., Epp, 2003; Durand-Guerrier et al., 2012) have suggested that the transition to the acceptance of the principles that rule the understanding of mathematical statements could be supported by the use of examples in a familiar context whose interpretations agree with its mathematical interpretation. Nonetheless, the case of Andrea is interesting because it was a bit different. She understood that "some-statements" *may* be universally true (which automatically rejected her initial assumption dAA7[3]) with the support of a mathematical context example, and then she transferred that understanding to the familiar-context examples she produced. Furthermore, the number of cases involved in Andrea's examples are finite, in contrast to the mathematical conflicting example that involved an infinite number of cases.

An additional important feature is the nature of the truth of the two universal statements involved in Andrea's examples. Statements St93 and St96 had an interesting and differing characteristic in respect to the sufficient evidence required to guarantee their truth.

St93: *All got an "A" in a test*

St96: *All teachers in the classroom are women*

[158] For details on these types of responses, see Chapter 3, Section I.1.2.

Andrea claimed that St93 and St96 were true and from there she referred to the truth value of the respective “some-statement”; however, the basis for the truth of the universal statements relied on two different processes. Specifically, while the truth of St96 could be easily verifiable since each teacher in the classroom at that moment was indeed a woman, the truth of St93 relied on a demand that Andrea made when she asked us to *assume* that St93 was true. This is similar to the case of *imaginary* statements the teachers were asked to analyze during and after the intervention, where the teachers were asked to assume the truth of some statements whose truth value was impossible to determine.

The truth value as a factor that might have influenced Andrea’s change of assumptions

After Discussion 7.4.3 we revisited a statement that was similar to one the teachers had seen in the First Exploratory Interview before the intervention. We discussed the imaginary statement St100,

> St100: *Some even numbers are Vallejo numbers.*

I asked the teachers to determine whether based on St100 it could be inferred that *some even numbers were not Vallejo numbers*. Andrea immediately showed her concern about the truth value of St100.

> Andrea: *But in that example I should be clear what the antecedent and the consequent are, because in the given statement it is not clear what Vallejo numbers are… and when we talk about some, I mean, how am I supposed to know whether this is true or false?*

Even though what Andrea expressed was right about the need to be clear what “Vallejo numbers” were in order to determine the truth value of St100, the task did not really require knowing its truth value in order to answer the question I posed.

Next, I highlighted the goal of the task and the teachers were requested to think about the following structure (that I wrote on the whiteboard).

Some even numbers are Vallejo numbers $\xrightarrow{?}$ Some even numbers are not Vallejo numbers

Since I wanted to check whether the truth value changed Andrea’s answer, I now asked the teachers to assume that the existential statement on the left was true. I asked whether the “some-statement” on the right could be inferred from St100, on the left. Andrea disagreed with it.

This short episode suggested that Andrea had an expectation to reason with true “some-statements” in the context of her new set of assumptions; however, as I show next, this was not precisely the case.

New understanding in action: Supporting a colleague’s understanding

After the intervention for teachers, Andrea managed to put into practice her new insights related to “some-statements”. Notably, during our Meeting #11 the teachers engaged in a discussion on whether the imaginary statement St141,

> St141: *Some RAINBOW numbers are smaller than 3,*

implied that *some RAINBOW numbers were not smaller than 3*. Andrea gave a negative answer and she added that she remembered our discussion of a similar example during the intervention for teachers. Furthermore, she said that she recalled that it was a *mathematical* example. In contrast to Andrea's answer, Lizbeth gave an affirmative answer[159]. In order to support Lizbeth's understanding, Andrea used one of the examples she came up with during the intervention. The example was framed in our current context (again we all were women in the classroom).

> Andrea: *No [St141 does not imply that some RAINBOW numbers are smaller than 3], because for example, I say, here, some people present in here are women. It does not mean that there are some who are not [women]... Even though all of us are [women], "some" can be used. I mean, even though here we all are women, I can also use some people in here are women.*

Andrea's contribution highlighted that despite the fact that the universal statement *"all people present in here are women"* was true, it was possible to make the claim that *some people present in there were women*. And certainly, the latter did not necessarily imply that *some people were not women*.

This shows that Andrea used her new insights to reason with imaginary statements like St141. Based on her new assumptions, she concluded that the affirmative "some-statement" St141 did not imply the respective negative "some-statement". The fact that she applied her new assumptions to the case of an imaginary statements suggests that her assumptions were not attached to specific truth values given that imaginary statements have no truth values.

Factors that influenced Andrea's responses to the conflict

Some factors played an important role in Andrea's responses to the conflict she faced during Discussion 7.4.2: (1) Andrea was familiarized with the statements included in the conflicting example. A crucial factor was that the two statements included in the conflicting example were within the reach of Andrea's mathematical background knowledge. The statements involved simple mathematical concepts like *even numbers* and *divisibility in the context of natural numbers*, which were covered in the first part of the intervention[160]. Here it was particularly important that she knew that all numbers divisible by 4 are even. In this sense, this first factor is similar to the *individuals' prior knowledge factor* that Chinn and Brewer's (1993) pointed out may influence how individuals respond to anomalous data[161]. (2) Andrea used the mathematical meaning of "some" as *at least one* when discussing the conflicting example. This factor was key to her acknowledgement that St26 was true and St27 was not, which was needed to become aware of the conflict.

> St26: *Some numbers divisible by 4 are even numbers*
>
> St27: *Some numbers divisible by 4 are not even numbers*

The teachers were given an initial input for the mathematical meaning of "some" ("some" means "at least one") and were asked to use this meaning as the intervention had a focus on developing *mathematical* understandings. Notably, this meaning allowed Andrea to conclude that St26 was true, while St27 could not be true, which was crucial to her

[159] Remember that Lizbeth was not present during Discussion 7, when we engaged on discussing existential statements.
[160] See Chapter 4, Section II.2.1 for details on the design of the intervention.
[161] For details on Chinn and Brewer's (1993) factors that may influence how individuals respond to anomalous data, see Chapter 2, Section III.1.

realization that her initial assumptions did not hold. (3) Andrea's attention was drawn to the main core of the conflict. The teachers' attention was guided to three aspects: the truth of St26, the fact that St26 was universally true, and the falsity of St27. In that respect, direct questions supported her explicit awareness (e.g., Is St26 true? Does St26 hold for all the elements in the set of analysis? Is St27 true?). This was a crucial factor for Andrea to notice the conflict in the first place. (4) Andrea needed to explain the conflict. The expectation for explanations was a *sociomathematical norm* (Yackel & Cobb, 1996) that I established from the onset of the intervention. It is one feature of Proof-based Teaching (Reid & Vallejo-Vargas, 2017), which is the context I used to develop the intervention. Such an expectation might have triggered Andrea's need to explain the conflict. Andrea formulated and illustrated her new assumptions as a way to do so. She identified that because there may be cases of true "some-statements" that were universally true, her initial assumptions did not longer hold. She used this observation to explain and modify her set of initial assumptions.

Summary of Section II.1

Andrea began the intervention for teachers with several initial assumptions about existential statements. Among them, Andrea made a distinction between "some-statements" and "there-is-statements". In her view, *"Some X are Y"* and *"There is X that is Y"* did not make the same claim; that is, the form of the existential statements made a difference in Andrea's initial interpretation. Evidence of this was Andrea's initial assumption that from a true US *"All X are Y"*, it can be inferred that *"There is X that is Y"* is true (her initial assumption bAA7) and that it will follow that *"Some X are Y"* is false (her initial assumption dAA1[1]). Andrea's initial meanings for the existential quantifiers "some" and "there is" had the daily-life context as its main source. While she associated the quantifier "there is" with *existence*, her interpretation of the quantifier "some" partitioned the set of analysis (*X*) into two parts (one that held *Y*, and the other that did not, when *"Some X are Y"* was considered). Her initial use of "some" as "from all, one group, but not all" (her initial assumption dAA1[8]) led to other two assumptions she used: *"All X are Y"* and *"Some X are Y"* cannot both be true (her initial assumption dAA1[1]) and *"Some X are Y"* implies that *"Some X are not Y"* (her initial assumption dAA7[3]).

During the intervention the introduction of the mathematical meaning for "some" (as "at least one") did not have an immediate impact in Andrea's initial related assumptions. As she was prompted to use this meaning, she became aware that one confirming example was sufficient to prove that a "some-statement" was true. Yet, her new mathematical insight on what established truth of a "some-statement" and its proving did not disturb her other initial assumptions dAA7[3] and dAA1[1]. Andrea needed to face a cognitive conflict in order to change those two assumptions. The conflicting example consisted of a true mathematical "some-statement" that was universally true. Andrea's responses to the cognitive conflict can be summed up as follows: (1) Noticing: Andrea noticed the main characteristics that made the given example a conflicting example with respect to what she previously assumed, which was according to a familiar-context framework. She noticed the implications of using a mathematical framework. (2) Generalizing: Andrea formulated a generic claim based on her observations. This was a result of identifying the most important characteristics in the "noticing" phase. (3) Contrasting: Andrea explicitly pointed out the way she used the quantifier *"some"* (*"some"* means *"some and not all"*) to contrast it with her new insights about its use in mathematics (*"some"* means *"at least one and maybe all"*). (4) Producing: Andrea produced two new examples that satisfied the same characteristics of the conflicting example, which she used to illustrate her new understanding. (5) Using her new insight over time: Andrea used one of the examples she produced during the intervention in a meeting after the intervention to support one of her colleagues' understanding of existential statements.

Andrea's responses show that she modified her initial assumptions. Her adjudgment of the interpretation of "some" to the mathematical context and the conflicting example during the intervention played a crucial role in Andrea's emerging understanding of existential statements.

Figure 33 shows the development of Andrea's assumptions about establishing truth and proving ESs and the elements or aspects of the intervention that might have supported or triggered changes[162].

[162] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

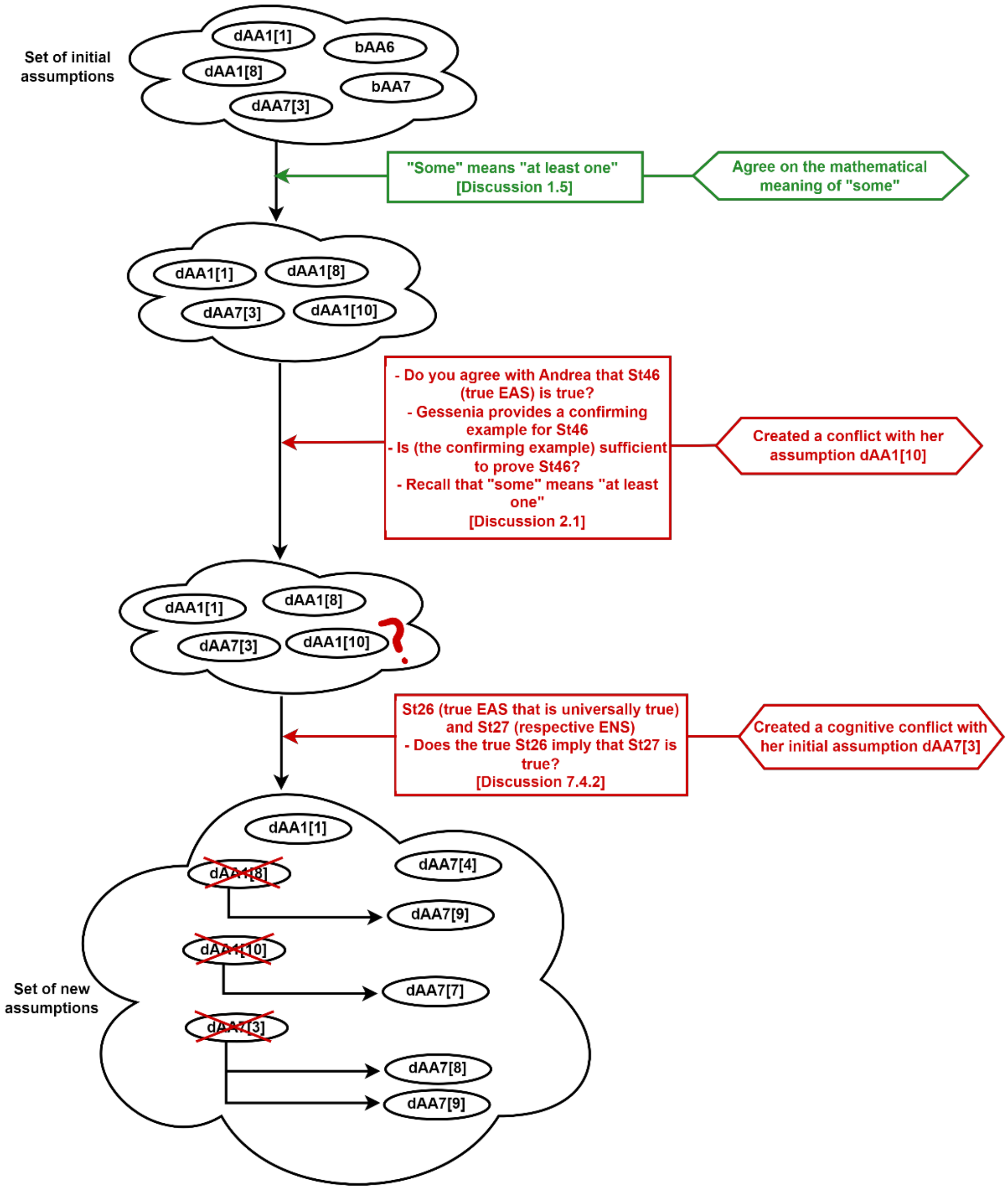


Figure 33. The development of Andrea's assumptions about establishing truth and proving ESs (bAA6: "There is X that is Y" does not convey quantity; bAA7: The true US "All X are Y" implies that "There is X that is Y" is true; dAA1[1]: The true US "All X are Y" implies that "Some X are Y" is false; dAA1[8]: "Some" means "from all, one group, but not all"; dAA7[3]: "Some X are Y" necessarily implies that "Some X are not Y"; dAA1[10]: Confirming examples are insufficient to prove a statement; dAA7[4]: One confirming example is sufficient to prove a "some-statement"; dAA7[7]: "Some" means "at least one, and maybe all"; dAA7[8]: If "Some X are Y" and "All X are Y" are both true, then "Some X are not Y" is false; dAA7[9]: If "Some X are Y" is true and "All X are Y" is false, then "Some X are not Y" is true)

2. Establishing Falsity and Disproving Existential Statements

From a mathematical point of view, an existential affirmative statement (EAS), for example the affirmative "some-statement" *"Some X are Y"*, is false if and only if *no X is Y*. Likewise, an existential negative statement (ENS), for example the statement *"Some X are not Y"*, is false if and only if *all X* are *Y*.

In this section I focus on the development of Andrea's understanding to establish falsity and disprove "some-statements" during the intervention. Her development went through a number of changes. I pay close attention to three different criteria Andrea used in this process. The first criterion was Andrea's initial assumption to establish falsity of "some-statements" and was directly linked to the initial interpretation she used for the quantifier "some". The second and third criteria she came up with during the intervention and were associated to the concepts and insights she developed in this process.

Table 17 includes the criteria Andrea used to establish falsity and disprove "some-statements" before, during or after the intervention. The assumptions in the table are presented as they emerged. The table includes the name I used for the approaches, the focus (whether the assumption was about an affirmative or negative "some-statement"), when it emerged (whether it was before, during or after the intervention), the basis for the approaches (the grounds that led Andrea to coming up with the assumptions) that includes its nature (whether it is mathematical or not).

Table 17. Andrea's approaches and assumptions to establish the falsity of "some-statements": their focus and basis

Approach	Assumption	Focus	When	Basis
(1) The "some = not all" approach	dAA1[1]: *"Some X are Y"* is false if *"All X are Y"* is true	Affirmative "some-statement"	Before the intervention	Non-mathematical nature: The meaning of "some" as "not all".
(2) The "counterexample" existence approach	dAA7[5a]: *"Some X are Y"* is false if it has a "counterexample"	Affirmative "some-statement"	During the intervention	Non-mathematical nature: Reasoning by analogy.
	dAA7[5b]: *"Some X are not Y"* is false if it has a "counterexample"	Negative "some-statement"	During the intervention	
(3) The "impossible to be true" approach	dAA7[6]: *"Some X are not Y"* is false if it is impossible to find an example of *X* that is not *Y*	Negative "some-statement"	During the intervention	Meaning of the "some-statement", which is based on the mathematical meaning of "some" as "at least one". The nature of the grounds will depend on the mathematical background knowledge she uses to guarantee the impossibility to find confirming examples for the statement.
	aAAt1: *"Some X are Y"* is false if it is impossible to find an example of *X* that is *Y*	Affirmative "some-statement"	After the intervention	

Next, I describe each of Andrea's approaches to establish the falsity of "some-statements", what they consist of, how and when they emerged as well as their grounds.

Andrea's initial approach: The "some = not all" approach

Andrea's initial assumption for establishing and proving the falsity of "some-statements" consisted in considering that a "some-statement" was false if it was universally true.

Andrea's initial[163] criterion to establish falsity and disprove an affirmative "some-statement" entirely relied on the truth value of the respective universal affirmative statement[164]. This was a direct consequence of the initial everyday meaning Andrea used for "some" ("some, but not all"). I have already discussed this in the previous section (see Section II.1 above); however, I include it here as well because it revealed Andrea's criterion to establish falsity. The main point is that Andrea's initial meaning for "some" led her to reject the possibility that the two statements *"Some X are Y"* and *"All X are Y"* could be both true. Based on Andrea's initial assumption, once one of the statements, *"Some X are Y"* or *"All X are Y"*, was determined to be true, the other, as a direct consequence, was false: *"Some X are Y"* is false if *"All X are Y"* is true (her assumption dAA1[1] in Table 16). I called her approach the "some = not all" approach since it relied on her interpretation of "some" as "not all".

For example, as I pointed out in the previous section, right before Discussion 1, she used this criterion to determine that the existential statement St26:

> St26: *Some numbers divisible by 4 are even numbers*

was false as she was aware that it was universally true (i.e., she was aware that *"All numbers divisible by 4 are even"* was true).

Given that Andrea based her criterion to establish the falsity of a "some-statement" on the meaning of "some" from the daily-life context, its nature is non-mathematical.

Andrea's second approach: The "counterexample" approach

Andrea's initial assumption experienced its first change during Discussion 7. Andrea used the mathematical scheme she had previously built to disprove UASs (*counterexamples disprove UASs*, see Section I.2.1.2 above) and extended it to the case of existential statements.

During Discussion 7.1 Andrea argued that the "some-statement" St24 was false[165].

> St24: *Some divisions of natural numbers are exact divisions*

She claimed that she could give a "counterexample" for St24.

> Andrea: *This [St24] would be false because I can pick one [example] that doesn't hold… because at least one division does not satisfy that [being exact].*

[163] By "initial criteria" I mean the criteria Andrea started with, for which the intervention had no apparent direct influence.

[164] I call the "respective" UAS of the statement *"Some X are Y"* to the statement *"All X are Y"*.

[165] For additional details on this discussion, see Section II.1 above.

Andrea needed to come up with a mathematical strategy to disprove "some-statements", since as she pointed out we had not discussed any false "some-statements" yet (*"Well, the thing is that we haven't seen, um, this sort of cases"*). So, she extended her understanding of disproving UASs and assumed that *"counterexamples" could also disprove "some-statements"* and concluded that St24 would be false. Moreover, Andrea used her insights related to the characterization of counterexamples to a UAS (*a counterexample must satisfy the antecedent but not the consequent of the statement*, see Section I.2.2.2 above) to describe a "counterexample" for St24 (*"because at least one division does not satisfy that [being exact]"*). An abstract version of Andrea's current assumption and her presumed rationale is shown in Figure 34.

> *"All X are Y"* is false because I can find a counterexample for it (an *X* that is not *Y*).
> Similarly, *"Some X are Y"* would be false because it has a "counterexample" for it (an *X* that is not *Y*).

Figure 34. Andrea's assumption dAA7[5a] for the falsity of an affirmative "some-statement" and its disproving.

In Piaget's terms, Andrea explained her disproving by *assimilating* the case of affirmative "some-statements" to her disproving scheme for UASs to have a mathematical approach to disprove "some-statements".

The main issue with Andrea's new assumption was that it undermined the consistency of her current set of related assumptions; however, she did not seem aware of this. Observe that because of her still current use of her initial meaning for "some" ("some, but not all"), *"Some X are Y"* (here St24) was true, which was already a contradiction with her new assumption that it was false due to the existence of a "counterexample".

Her "counterexample" approach was indirectly rejected by Gessenia, who showed that St24 was true[166]. However, Andrea used her "counterexample" approach after Discussion 7.1, to disprove a different sort of "some-statement": a *negative* "some-statement".

After Discussion 7.1 Gessenia gave the universal affirmative St86 and its respective negation St87[167].

St86: *All natural numbers that end in digit 0 are multiples of 5*

St87: *Some natural numbers that end in digit 0 are not multiples of 5*

Andrea showed interest in learning how to disprove the "some-statement" St87. Andrea assumed that a counterexample would disprove St87.

Andrea: *If it is false, then I would give a counterexample.*

Andrea's third assumption is a manifestation of her "counterexample" approach to disprove negative "some-statements" (see Figure 35).

> The statement *"Some X are not Y"* is false because it has a "counterexample" for it (presumably, an *X* that is *Y*).

Figure 35. Andrea's dAA7[5b] assumption for the falsity of negative "some-statements" and its disproving.

[166] See Section II.1 for details above.
[167] For details on this discussion, see Section I.4, above.

Presumably, Andrea was motivated by three reasons that supported her new extension of the "counterexample" approach to the case of negative "some-statements": first, Andrea was aware that St87 was false. Her mathematical background knowledge made her aware that St86 was true. Second, she was aware that there were potential "counterexamples" to St87: an example that satisfied the first condition of the statement (i.e., a number that ended in digit 0) and that contradicted the second condition (i.e., that was a multiple of 5); for example, the number 20. Third, a presumption that the "counterexample" approach had not been discarded for the case of negative "some-statements", but only for the case of affirmative "some-statements" (see Figure 36).

> *"All X are Y"* is false because I can find a counterexample (an *X* that is not *Y*) for it.
>
> The statement *"Some X are Y"* would be false because it has a "counterexample" for it.
>
> *"Some X are Y"* is actually true.
>
> Then the "counterexample" approach does not work for affirmative "some-statements".
>
> I am certain that the negative "some-statement" *S* (S: *Some X are not Y*) is false and I am sure that there is an *X* that is *Y*.
>
> The counterexample approach did not work for affirmative "some-statements", but it might still work for negative "some-statements".
>
> *S* must be false because it has a "counterexample" (an *X* that is *Y*) for it.

Figure 36. Andrea's updated rationale for her approach to determine the falsity of "some-statements"

As in the case of affirmative "some-statements", Andrea *assimilated* the case of negative "some-statements" to her disproving scheme for UASs due to the three motives I gave above. For Andrea, providing a "counterexample" seemed to be a reasonable justification.

Neither of Andrea's assumptions were aligned with the way a "some-statement" is proved to be false in mathematics. Andrea's rationale for her new assumptions was non-mathematical.

Andrea's "counterexample" approach has been also shown in previous research with in-service elementary school teachers (Barkai et al., 2002) and school students (Buchbinder & Zaslavsky, 2009). In those studies, though, the existential statements used were affirmative "there-exist-" and "there-is/are-" statements. It is not clear whether the same individuals would have used the "counterexample" approach to disprove affirmative "some-statements" or negative "there-exist-", "there is-" and "some-" statements[168].

Andrea's third approach: The "impossible to be true" approach

Andrea's change of her second approach to disprove "some-statements" was the result of two factors: Andrea's awareness of what it meant for a "some-statement" to be true (see Section II.1 above) and her background knowledge about divisibility (the mathematical content involved).

Andrea's third approach to establish falsity of and disprove "some-statements" emerged as we continued the discussion in the previous section, after Discussion 7.1. Episode 16 shows the dialogue we engaged in as a result of Andrea's interest in learning how to disprove the "some-statement" St87, which resulted after negating St86.

St86: *All natural numbers that end in digit 0 are multiples of 5*

[168] For details on research related to this issue, see Chapter 2, Section I.7.2.

St87: *Some natural numbers that end in digit 0 are not multiples of 5*

Episode 16

Turn	Who	What
1	Andrea	*How can I disprove the second statement [St87]?*
		…
2	Andrea	*If it is false, then I would provide a counterexample.*
3	I	*No*
4	Andrea	*What then?*
5	I	*Its negation is a universal statement that is true. So, its justification needs to be general and involve an infinite number of cases.*
6	Andrea	*[---]*
7	I	*I mean, we know that if we have an existential statement that is true, I should show at least one case that holds the statement, right?*
8	*[The three teachers nodded in agreement]*	
9	I	*At least one; however, if it is false, one counterexample is not sufficient. A counterexample is sufficient when the statement is universal and false.*
10	Andrea	*Right.*
11	I	*But when it is existential and false, it [disproving it] is the same as proving that its negation is true; that is, that this one [St86] is true.*
12	Andrea	*Ahhhh!!!*
13	I	*It means that this [disproving St87] turns into proving a true universal statement. And here [St86] there are infinite cases, then, that is why I told you, we would need a general justification.*
14	Andrea	*And its justification [for St87], then here-*
15	I	*Relies on the justification of its negation [St86]. This one (I pointed to St86).*
16	Andrea	*Can't I pick examples?*
17	I	*No*
18	Andrea	*Because I am not going to find any examples!*

I rejected Andrea's "counterexample" approach to disprove the statement St87 (turn 3). I drew the teachers' attention to the negation of St87 (turn 5), what was involved in proving an ES (turn 7), the sufficiency of one counterexample for the case of universal statements, but not for the case of existential statements (turn 9). I explained that disproving St87 was the same as proving its negation[169] St86 (turn 11) and since it involved infinite cases, a general justification was expected (turn 13). Andrea directed her attention to the possibility to provide confirming examples for St87 (turn 16). This I presume based on her conclusion that there would not be any of them (turn 18), given that it was indeed impossible to find numbers that ended in zero and were not multiples of 5. That is, she explained the impossibility for the "some-statement" to be true by relying on her background mathematical content knowledge (her assumption dAA7[6]). This I called Andrea's "impossible to be true" approach. In Figure 37 I include an abstract schema of Andrea's new form of reasoning.

Is it possible that *"Some X are not Y"* is true? That is, can I find *at least* one example of *X* that is not *Y*?

No, it is impossible (based on her background knowledge about the mathematical content).

Figure 37. Andrea's reasoning for her "impossible to be true" approach to disprove a negative "some-statement".

[169] For details on the negation of ESs, see Section II.3, below.

Andrea's reasoning implicitly involved proving that *"No X is not Y"* is true (i.e., that it is not possible to find at least one *X* that was not *Y*), which is equivalent to proving that *"All X are Y"* is true. In the case of the specific statement in discussion, disproving St87 involved proving that St86 (*"All natural numbers that end in digit 0 are multiples of 5"*) is true. Andrea used her background knowledge about the mathematical content involved (as she was aware that all natural numbers that ended in digit 0 were multiples of 5) to conclude that there would not be any such examples.

It is possible that my input to recall what was involved in proving an ES had drawn her attention to analyzing that possibility. This became clear as she used that form of reasoning later, during her teaching, as I will discuss next.

Traces of emergent understanding: Andrea's Teaching

Andrea's teaching of Session 7 included the analysis of the truth value of St131 and its respective justification.

> St131: *Some divisions by 4 have a remainder equal to 7*

Andrea explained as follows.

> Andrea: *Is there indeed a division like that? When I say "some", I mean "at least one". When I divide by 4, is the remainder going to be equal to 7? Let's see, why?... Here, were you asked for ALL divisions? ... It means, if there was at least one, if one holds Giussepe [calling for one student's attention], if one holds, this is valid. At least one division, it says, one division by 4, that has a remainder equal to 7. Let's see whether this is true or not… there will be some divisions, here it says, such that when you divide by 4, the remainder will be equal to 7? There will be some?...*

Andrea's feedback suggested her focus on three aspects when explaining how to establish falsity and disproving St131 with her "impossible to be true" approach. First, Andrea drew her students' attention to the mathematical meaning of "some" as "at least one" and she incorporated this meaning to the logical interpretation of St131 (*"When I say "some", I mean "at least one" ... At least one division, it says, one division by 4, that has a remainder equal to 7"*). Second, Andrea invited her class to think about the possibility for St131 to be true (*"there will be some divisions, here it says, such that when you divide by 4, the remainder will be equal to 7?"*). Third, her explanation for the impossibility that St131 could be true was clearly stated in the final conclusion she wrote down in front of her class, as it reads in Figure 38. Andrea based the impossibility for St131 to be true on a property that her class had previously proved about the maximal remainder in a division (the maximal remainder in a division by *"n"* is *"n-1"*).

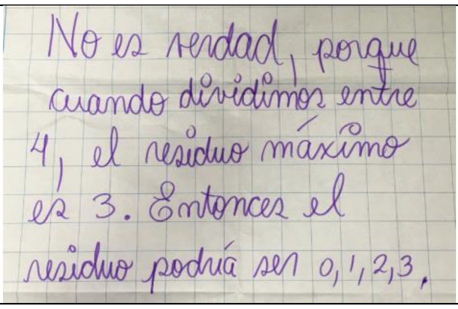
 No es verdad, porque cuando dividimos entre 4, el residuo máximo es 3. Entonces el residuo podría ser 0, 1, 2, 3.	*It is not true, because when we divide by 4, the maximal remainder is 3. So the remainder could be 0, 1, 2, 3.*

Figure 38. Andrea's class's conclusion for the truth value and justification of the statement "Some divisions by 4 have a remainder equal to 7". [The original Spanish is on the left side; while an English translation is on the right]

This episode is interesting because here Andrea used her "impossible to be true" approach for the case of an affirmative "some-statement" (her assumption aAAt1 in Table 17), which contrasted with her first use of the approach for the case of a negative "some-statement" during the intervention (her assumption dAA7[6] in Table 17).

Andrea's "impossible to be true" approach is similar to *the negation of the meaning* method to negate a single-level quantification introduced by Dubinsky, Elterman and Gong (1988). Their method *"consists in realizing that a universal (existential) quantification asserts that every (at least) one of a collection of propositions has the value true, so its negation is the assertion that at least (every) one of them is false"* (p. 48). It is interesting that, unlike Dubinsky and his colleagues' focus, the strategy Andrea used was focused on *disproving* "some-statements" and not on *negating* them. Even though there is a direct relation between those two processes, Andrea's attention was not yet on the negation of "some-statements" when her approach emerged.

Unlike Andrea's two first approaches, the "impossible to be true" strategy was close to a mathematical approach to disproving existential statements.

Summary of Section II.2

Andrea's assumptions about what was involved in establishing the falsity of and disproving a "some-statement" went through different changes during the intervention. Andrea revealed her initial assumption before the intervention. It relied on her initial interpretation of the existential quantifier "some" (as "some, but not all"). Andrea assumed that a statement of the form *"Some X are Y"* was false if the statement *"All X are Y"* was true (her assumption dAA1[1]). This I call Andrea's "some = but not all" approach.

Her second approach was observed during Discussion 7. It involved an extension of her disproving schema for UASs, first, to the case of an affirmative "some-statement". Andrea assumed that a statement of the form *"Some X are Y"* was false if there was a counterexample for it (presumably an *X* that was not *Y*) (her assumption dAA7[5a]). This I call Andrea's "counterexample" approach. Andrea's approach was indirectly rejected by showing that the affirmative "some-statement" was actually true. After that discussion Andrea used again her "counterexample" approach, though for the case of a negative "some-statement" (her assumption dAA7[5b]). This suggested that she presumed that even though her "counterexample" approach was not valid for the case of affirmative "some-statements", it may be valid for the case of negative ones. I explicitly rejected her approach. I drew the teachers' attention to the logical interpretation of the statement and what was involved in proving it. I attempted to make connections between the falsity of the negative "some-statement" and the truth of its negation. However, from all my inputs, what seems to have supported Andrea's understanding of the disproving of "some-statements" was the logical interpretation of the statement plus her observation that it was impossible that it could be true (her assumptions dAA7[6] and aAAt1). Her third approach I called Andrea's "impossible to be true" approach and it was observed again during her teaching.

Unlike Andrea's two first approaches, the third was close to a mathematical form of reasoning. However, whether its nature was mathematical or not depends on the type of evidence she might have used to guarantee why it is impossible to find confirming examples for the "some-statement" under discussion. For example, during her teaching she used a property her class had previously proved, which meant that the nature of the evidence involved in her approach to disprove the respective "some-statement" was mathematical.

Figure 39 shows the development of Andrea's assumptions related to establishing falsity and disproving ESs and the elements or aspects of the intervention that might have supported or triggered changes[170].

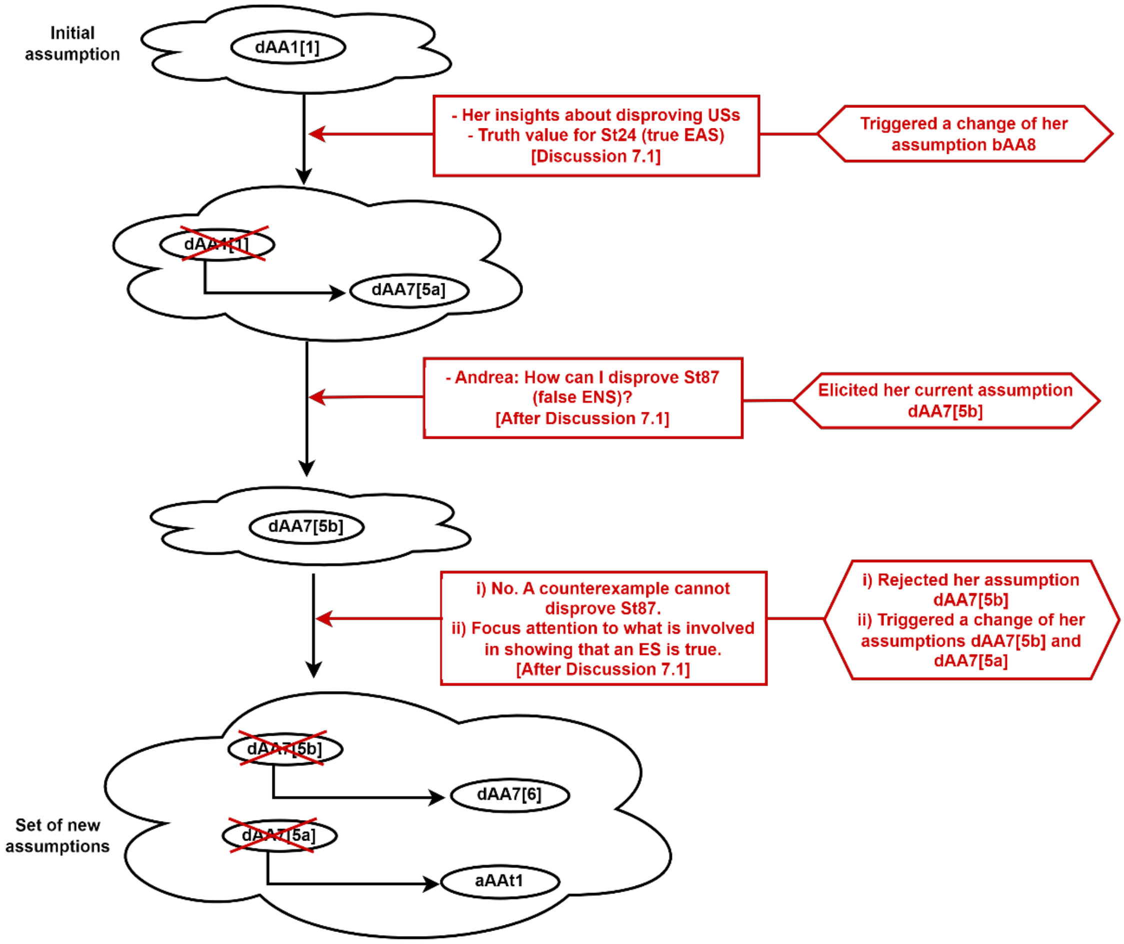


Figure 39. The development of Andrea's assumptions about establishing falsity of and disproving "some-statements" (dAA1[1]: "Some X are Y" is false if "All X are Y" is true; dAA7[5a]: "Some X are Y" is false if it has a "counterexample"; dAA7[5b]: "Some X are not Y" is false if it has a "counterexample"; dAA7[6]: "Some X are not Y" is false because it is impossible to find an example of X that is not; aAAt1: "Some X are Y" is false because it is impossible to find an example of X that is Y)

3. Negation of Existential Statements

The negation of existential statements came up in three contexts. The first episode arose after Discussion 7.1. The negation of a "some-statement" was used as a way to explain what was involved in disproving it; namely, I attempted to establish that disproving the "some-statement" entailed the same as proving its negation. Second, Discussion 7.5 included the negation of a universal statement that led to a discussion about the negation of "some-statements" and whether it produced "all-" or "no-" statements. The discussion was introduced by Andrea, who was clearly puzzled by this issue. Finally, Discussion 9 was the first opportunity where the teachers dealt with the negation of (one affirmative

[170] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

and one negative) “there-exist-statements”. The teachers were expected to determine whether the negations were existential or universal statements, their truth value and justification.

Table 18 includes Andrea’s assumptions related to the negation of existential statements as they were first observed before, during and after the intervention.

Table 18. Andrea’s assumptions about the negation of ESs

Assumption	**Type of assumption (when first observed)**
dAA9[2]: *“There does not exist X that is Y”* is false because there is indeed an *X* that is *Y* (i.e., because *“There exist X that is Y”* is true)	Initial assumption (before the intervention)
dAA9[3]: There does not exist X ≡ No X	Initial assumption (before the intervention)
dAA9[4]: *“There does not exist X that is Y”* is the same as *“No X is Y”*	Initial assumption (during the intervention)
dAA9[5]: *“There does not exist X that is not Y*” is false because there exist an *X* that is *Y.*	Initial assumption (during the intervention)
dAA9[6]: *“There does not exist X that is not Y”* is the same as *“No X is not Y”*	Initial assumption (during the intervention)
dAA7[15]: The negation of a *“some-statement”* is an *“all-statement”,* and never a *“no-statement”*	New assumption (during the intervention)
dAA7[16]: The negation of *“Some X are Y”* (*“Some X are not Y”*) is *“All X are not Y”* (*“All X are Y”*)	New assumption (during the intervention)
dAA7[21]: The negation of *“Some X are Y”* is *“No X is Y”*	New assumption (during the intervention)
dAA9[7]: *“There does not exist X that is not Y”* is false because there is an *X* that is not *Y*	New assumption (during the intervention)
dAA9[8]: *“There does not exist X that is not Y”* is the same as *“No X is Y”*	New assumption (during the intervention)

I split this section into two parts: negation of “some-statements” (Section 3.1), and negation of “there-exist-statements” (Section 3.2). I separated existential statements in these two groups (“some-” and “there-exist-” statements) because the form of the existential statements influenced their negation and related assumptions, as I began to show in Section II.1.

3.1. Negation of “some-statements”

In this section I focus on a conflict Andrea faced when dealing with “some-statements” and their negations. This conflict was evidence of Andrea’s uncertainty about what the negation of a “some-statement” was. The conflict revealed Andrea’s lack of awareness of the relation between “no-statements” and “all-statements”, but mainly the way this impacted her understanding of the negation of “some-statements”.

Discussion 7 had as a main focus the introduction of existential statements in relation to the negation of universal statements (see Section I.4). The negation of a “some-statement” was indirectly introduced after Discussion 7.1 (see Section II.2). After Discussion 7.1 the teachers were asked to provide a universal statement and negate it. Gessenia provided statement St86, and St87 as the negation for St86.

St86: *All natural numbers that end in digit 0 are multiples of 5*

St87: *Some natural numbers that end in digit 0 are not multiples of 5*

Andrea's attention was on the justification that disproved St87. Episode 16 (in Section II.2 above) includes the dialogue developed about this issue. From Episode 16 I focus here on two main observations in relation to the negation of a "some-statement": first, I took for granted that the teachers could find "evident" that the negation of St87 was St86 (see turns 5 and 11); however, as I found out after this episode that Andrea was not sure whether the negation of a "some-statement" was a universal statement (see *The Conflict* below). Second, Andrea's attention was not completely on the negation of St87, and moreover, on the link between disproving St87 and proving its negation St86. This was suggested by Andrea's surprised reaction (turn 12), which revealed her lack of awareness about this linkage.

Episode 16 shows the implicit way I used to introduce the negation of existential statements (turns 5 and 11). This can be expressed as in Figure 40.

If the negation of a universal affirmative statement UAS is the existential negative statement ENS, then the negation of the existential statement ENS is the universal statement UAS, from which it was obtained in the first place.

Figure 40. An implicit introduction to the negation of Existential Statements.

Furthermore, I described the conditions needed to disprove St87 in terms of proving St86 (see Figure 41). In the process of showing that disproving St87 was the same as proving St86, I ignored that the teachers had not noticed in the first place that the negation of St87 was St86. I naively assumed that the teachers had grasped that because *"Some X are not Y"* was the negation of *"All X are Y"*, then the negation of *"Some X are not Y"* would be *"All X are Y"* (Figure 40); however, I think this did not happen.

The negation of St86 is St87, then Disproving St87 = Proving the negation of St87 = Proving St86

Figure 41. The link between disproving St87 and proving St86.

At the end of this episode two things were clear about Andrea's awareness about the negation of ESs: (1) Andrea did not become aware of my introduction of negation of "some-statements" (Figure 40) and (2) she did not reason about the falsity of St87 by relying on my explanation that linked it to its negation (Figure 41). Instead, for the latter Andrea resorted to what I called her "impossible to be true" approach, which I explained in Section II.2.

The Conflict: If I negate "some", does it become "all" or "none[171]"?

During Discussion 7.5 we analyzed the simple implicit negation of a universal statement, St103:

St103: *Not all natural numbers divisible by 2 are divisible by 3.*

Here I focus on the discussion we had about its truth value and justification. Andrea relied on its equivalent statement St106:

St106: *Some natural numbers divisible by 2 are not divisible by 3.*

[171] Here I use the quantifiers "none" and "no" interchangeably. Hence, I sometimes use "none-statements" instead of "no-statements".

After Discussion 7.5, as a way to summarize the discussions we had, I prompted the teachers to consider *implicit negations* (like the "not" in St103) as affecting the whole quantified statement structure. In terms of the cases we had seen, this involved changing the universal for an existential quantifier and negating the second condition in the statement (see Section I.4 above). I concluded my feedback by telling the teachers that a similar reasoning would also apply for the case of existential statements, which triggered a discussion initiated by Andrea. It showed Andrea's interest in learning what needed to be negated in order to get a "no-statement". Episode 17 includes part of this discussion.

Episode 17

Turn	Who	What
1	I	*Even if I have the negation of an existential statement, it [the negation] is going to affect also both, my antecedent and consequent.*
2	Andrea	*If I negate an existential statement, does it become universal?*
3	I	*That's right.*
4	Andrea	*But, a positive or negative universal [statement]?*
5	I	*It depends on whether your original existential statement was affirmative or negative, doesn't it?*
6	Andrea	*No, but "some" is not positive, nor is it negative.*
7	I	*Some (I point to St106 that was written on the whiteboard), ta ta tan (expression I used to avoid reading "natural numbers divisible by 2" and emphasize other parts instead), are not-*
8	Andrea	*No, no, no.*
9	I	*And some are-, that is to what I refer.*
10	Andrea	*I mean, if I negate that "some", does it become "all" or "none"?*
11	I	*If you negate this one (I point to St106)? What happens if we negate this one [St106]?*
12	Andrea	*It is going to become "all".*
13	I	*How is it going to change? All numbers-*
14	Gessenia	*Divisible by 2 are divisible by 3.*
15	Andrea	*Alright, then it will never become "none".*
16	I	*"None", we will talk about "none" in short, because "none-statements" have a translation, or an equivalent-, they are universal. I think "none-statements" come next.*

My observation that the negation of an existential statement affected the antecedent and consequent of the statement (turn 1), led Andrea to asking (which revealed her continuing uncertainty about) whether the negation of an existential statement was a universal statement (turn 2). My confirmation (turn 3) resulted in Andrea's new request for precision on what kind of universal statement was expected; namely, a positive or a negative universal statement (turn 4). Based on what followed in the discussion, I interpret that by "positive universal statement" Andrea meant an affirmative "all-statement"; whereas, by "negative universal statement" she meant an affirmative "none-statement". My interpretation is supported later by turn 10. Andrea attributed a particular characteristic to the quantifiers "all" and "none": they were "positive" and "negative", respectively. Andrea was not paying attention to what I was paying attention to, which was the second clause of the statement, and whether it was in its affirmative or negative form (turns 5, 7 and 9). Further evidence of this is Andrea's response that "some" was not affirmative, nor was it negative (turn 6), which clearly revealed her exclusive focus on the quantifiers. In contrast to her initial assumption that "not all" was the same as "some" (see Section I.4), she did not count on a *semantic substitution* (Dawkins & Cook, 2017; Dawkins, 2017) for "not some" (turn 10). This might be explained by the everyday-language negation of a "some-statement". For example, when negating a "some-statement" of the form *"Some X are (not) Y"*, we normally add the expression *"It is not the case that"*, *"It is false that"*, or *"It is not true that"* right to the front of the "some-statement" (e.g., *"It is not the case that some X are (not) Y"*). This means that in such a context there is no way to simply place the negator "not" in front of the "some-statement", in contrast to what is possible with "all-statements", and also with existential statements of the form *"There exists X that is (not) Y"* and *"There is X that is (not) Y"*. In other words, there are no *simple implicit negations*, as I call them, for "some-statements". In Section 3.2 (below) I return to this point. Here it is important to emphasize that lacking an available substitution for "not some" might have been an obstacle for Andrea's understanding of the negation of "some-statements".

Furthermore, Andrea was not aware that "none-" and "all-" statements are directly linked by equivalences that are possible between them[172]. This was shown, for example, when I requested the teachers to negate St106 (*Some natural numbers divisible by 2 are not divisible by 3*); they hesitantly (though correctly) obtained St107.

St107: *All natural numbers divisible by 2 are divisible by 3.*

Andrea rushed to conclude that when negating "some-statements", the resulting negation would be an "all-statement", but never a "no-statement" (turn 15) (her new assumption dAA7[15]).

Andrea's lack of awareness of the relation between "all-" and "no-" statements was exhibited again later with her use of her assumption dAA7[15] as we discussed the "no-statement" St108,

St108: *No odd number is an Innova number.*

My approach to supporting Andrea's realization of what to negate in order to get a "no-statement" consisted of making her aware of the equivalence between St108 and St109,

St109: *All odd numbers are not Innova numbers.*

I expected that she would eventually understand that by negating *"Some odd numbers are Innova numbers"*, she would obtain the "no-statement" St108 through its equivalent "all-statement" St109.

Even though we came to a point where Andrea agreed[173] that St108 was equivalent to St109, Andrea kept asking the same question (what should be negated in order to get a "no-statement"?). This was a clear sign that it was difficult for Andrea to establish a meaningful link between a "no-statement" and its equivalent "all-statement". As I show in Section III.1 (below), the main source of Andrea's difficulties understanding the negation of "some-statements" was her own initial non-mathematical interpretation of the quantifier "no". Andrea's initial interpretation of "no X" prevented her from "seeing" the connection between statements of the form *"No X is Y"* and *"All X are not Y"*. As a result, and at least at this stage, she refused to accept that the negation of a "some-statement" could be a "no-statement".

Throughout the discussion Andrea insisted on knowing what statement needed to be (directly) negated to obtain a "no-statement". With some struggles Andrea eventually

[172] For details on Andrea's initial assumptions related to her lack of awareness about the relation between "no-" and "all-" statements, see Section III below.

[173] "Agreed" in the sense that Andrea showed her explicit agreement that both statements were equivalent; however, she did not exhibit her awareness of this equivalence later, when needed. See Section III.1 below for more details.

accepted that the negation of a "some-statement" may be a "no-statement"; however, her agreement was not based on any form of understanding, but instead on the authority of my inputs[174] and a pattern she saw in the examples.

> Andrea: *I think I have found a pattern. When the "some" [statement] does not have a negation in the consequent, when I negate that "some" statement, it is going to become [a] "none" [statement].*

The pattern in Andrea's claim can be expressed as shown in Figure 42.

The negation of *"Some X are Y"* = *"No X is Y"*

Figure 42. Andrea's new assumption dAA7[21] for what to negate to obtain a "no-statement".

This shows that Andrea seemed to have at least formally accepted that by negating a "some-statement" a "no-statement" was obtained. Even though Andrea did not specify the type of "no-statement" her claim referred to, I presume that it was an affirmative "no-statement" (i.e., *"No X is Y"*). My presumption is based on the cases of "no-statements" we had analyzed until that moment (only affirmative ones), and the rarity of negative "no-statements" (i.e., *"No X is not Y"*) in natural language.

3.2. Negation of "there-exist-statements"

In this section I show that Andrea's initial approaches to finding an equivalent statement for and disproving the simple implicit negation of an *affirmative* "there-exist-statement" were aligned with the mathematical perspective. I explain that the form of the statements had a big influence on Andrea's initial assumptions. In addition, I show that the case of simple implicit negations of *negative* "there-exist-statements" triggered a conflict that uncovered assumptions that Andrea had made about negations during the intervention.

During the intervention the teachers were given two examples of what I call *simple implicit negations* of existential statements; that is, ESs for which it is easy to incorporate the negator *"not"* in the statement to express its negation. For example, an implicit negation for *"There is X that is Y"* could be *"It is false that there is X that is Y"*, and a simple implicit negation for the same statement is *"There is not X that is Y"*. In this section, the forms of the simple implicit negations we analyzed were: *"There does not exist X that is Y"* and *"There does not exist X that is not Y".*

From a mathematical perspective, *"There does not exist X that is Y"* is equivalent to *"No X is Y"* and *"There does not exist X that is not Y"* is equivalent to *"All X are Y"* (or *"No X is not Y"*). In Section II.3.1 (above) I included the two ways I attempted to introduce the negation of ESs: after Discussion 7.1 I introduced the negation of a "some-statement" in an implicit way (see Figure 40), and after Discussion 7.5, I did it explicitly (*"if I have the negation of an existential statement, it [the negation] is going to affect also both, my antecedent and consequent"*, turn 1 in Episode 17 above).

In Discussion 9 the teachers were asked to solve Activity 13[175] that includes eight mathematical statements. The teachers were asked to group the statements with equivalent ones, determine whether the statements were universal or existential, true or false, and explain why. All the eight statements involved the same antecedent and

[174] For the details I omitted here, see Section III.1.3 below.
[175] For details on the content of Activity 13, see Appendix A13.

consequent. Among the eight statements two negations of “there-exist-statements” were given (numbered #5 and #8 in Activity 13).

> St123: *There do not exist divisions of natural numbers that have a remainder equal to 0*
>
> St126: *There do not exist divisions of natural numbers that do not have a remainder equal to 0*

St123 is the simple implicit negation of an affirmative “there-exist-statement” since the consequent is in its affirmative form, *“have a remainder equal to 0”*; whereas St126 is the simple implicit negation of a negative “there-exist-statement” given that its consequent is in its negative form, *“do not have a remainder equal to 0”*.

The fact that Discussion 9 included the first negations of “there-exist-statements” suggests that Andrea had mainly used her initial assumptions about them before then.

3.2.1. Simple implicit negation of affirmative “there-exist-statements”

There were two circumstances in which Andrea’s initial assumptions about the simple implicit negation of affirmative “there-exist-statements” were observed: her disproving of them and her determination of equivalent statements for them.

Disproving the simple implicit negation of an affirmative “there-exist-statement”

Andrea’s ease in disproving statement St123 (*“There do not exist divisions of natural numbers that have a remainder equal to 0”*) suggested that she used her initial approach to disprove implicit negations of affirmative “there-exist-statements”. She proved that statement St123 was false because of the division of 20 by 5, which shows that there exists a division of natural numbers that has a remainder equal to zero.

In general terms, Andrea’s initial approach to disproving statements of the form *“There does not exist X that is Y”* consisted of showing that *in fact there exists an X that is Y*. In other words, she disproved the original statement *“There does not exist X that is Y”* by proving *“There exist X that is Y”* (her initial assumption dAA9[2]).

Andrea’s initial assumption dAA9[2] is mathematically accurate; however, her set of assumptions about simple implicit negations of “there-exist-statements” as a whole was not consistent at that time, as I show next.

Equivalence for the simple implicit negation of an affirmative “there-exist-statement”

When asked to find an equivalent statement for the simple implicit negation of an affirmative “there-exist-statement” Andrea revealed her use of a *semantic substitution* (Dawkins & Cook, 2017; Dawkins, 2017). She substituted “there does not exist” with “none” (her initial assumption dAA9[3]) as part of her initial approach to find an equivalent statement for St123 (*“when I say “there does not exist” is “no”, for me, okay?”*, see the complete excerpt below in Section 3.2.2).

I asked the teachers to find an equivalent statement for St123 and to show step by step the transition from St123 to its equivalent “no-statement”, by specifying the intermediate equivalent “all-statement”. Specifically, I expected the teachers to identify that St123 (*“There do not exist divisions of natural numbers that have a remainder equal to 0”*) is equivalent to *"All divisions of natural numbers do not have a remainder equal to 0"*, and that this statement is equivalent to St125,

St125: *No division of natural numbers has a remainder equal to 0.*

Even though Andrea struggled to find an equivalent statement in this way[176], she did not reject the resulting statement St125. In fact, Andrea concluded that the statements St123 and St125 were equivalent for a different reason from the one I expected. She used her *semantic substitution* to "trade" the first part of St123 and obtain St125, as shown in Figure 43. In Section I.4 (above) I called this Andrea's *separate and substitute* (SS-) approach, which she used to find equivalent statements for the simple implicit negation of "all-statements", where she substituted "not all" for "some".

There does not exist X that is Y = **[There does not exist** X**]** that is Y = **[No** X**]** is Y = No X is Y

Figure 43. Andrea's initial "separate and substitute" approach to find an equivalent statement for "There does not exist X that is Y"

Here Andrea's initial assumption involved using her SS-approach to the case of simple implicit negations of "there-exist-statements". This encompassed her access to a semantic substitution for "there does not exist", which she traded with "no". As a result, she used her initial assumption that *"There does not exist X that is Y"* is the same as *"No X is Y"* (assumption dAA9[4]).

3.2.2. Simple Implicit negation of negative "there-exist-statements": A conflict

As in Section 3.2.1, the disproving of and the determination of an equivalent statement for a simple implicit negation of a *negative* "there-exist-statement" gave access to Andrea's assumptions related to the negation of negative "there-exist-statements".

Disproving the simple implicit negation of a negative "there-exist-statement"

In order to disprove St126 (*"There do not exist divisions of natural numbers that do not have a remainder equal to 0"*) Andrea explicitly showed her agreement with Gessenia, who provided an example of a division that had a remainder equal to 0.

Gessenia: *20 divided by 4… we see that there does exist… 20 divided by 4 is a division of natural numbers that shows that there do exist natural numbers that [when divided] give a remainder equal to 0. That is a counterexample.*

Generally speaking, Andrea agreed that *"There does not exist X that is not Y"* was false because there existed an *X* that was *Y* (her initial assumption dAA9[5]).

In contrast to her approach to disprove simple implicit negations of *affirmative* "there-exist-statements", Andrea's apparent initial assumption for the case of negative "there-exist-statements" was not mathematically correct.

Equivalence for the simple implicit negation of a negative "there-exist-statement"

I rejected Gessenia's disproving of St126 (*"There do not exist divisions of natural numbers that do not have a remainder equal to 0"*). Andrea's response can be

[176] For details on this issue, see Section III.1 below about "no-statements".

summarized in three steps: First, she determined an equivalent "all-statement" for St126, that is St43,

St43: *All divisions of natural numbers have a remainder equal to 0.*

Second, she determined that St43 is false and disproved it by relying on her current assumptions about false UASs (showing a counterexample is sufficient to disprove a US[177]).

Third, she concluded that St126 is false because its equivalent statement St43 is false. Andrea's new approach to determining the truth value of a statement of the form *"There*

> *"There does not exist X that is not Y"* is equivalent to *"All X are Y"*.
> The UAS *"All X are Y"* is false because there is an *X* that is not *Y* (a counterexample).
> Two equivalent statements have the same truth value.
> Then *"There does not exist X that is not Y"* is false.

Figure 44. Andrea's new evaluation of the truth value of the statement "There does not exist X that is not Y".

does not exist X that is not Y" can be expressed in abstract terms as in Figure 44.

In terms of the criteria Andrea used to find an equivalent statement for St126, I contend that Andrea first used an extension of her SS-approach (see Figure 43 above) to the case of simple implicit negations of *negative* "there-exist-statements" and obtained St127,

St127: *No division of natural numbers does not have a remainder equal to 0*.

She then matched St127 with her equivalent St43[178]. This became clear as the discussion continued and Andrea showed that she was aware of the equivalence among St126, St127 and St43 through her focus on St127.

> Andrea: *That [St127] is a weird example, isn't it?... that is an example of a weird case, because it is "no" and "do not have".*

In fact, I expected that Andrea would use her *separate and substitute* approach to find an equivalent statement for the implicit negation of the negative "there-exist-statement" St126. She already had a substitution for "there does not exist".

The Conflict

Here I show that Andrea experienced a conflict when she compared the negations she obtained using two different approaches.

After determining the truth value and justification for St126, Andrea experienced a cognitive conflict. It was triggered by contrasting what she obtained when using (1) her *separate and substitute* (SS-) approach (St127) with (2) what she would have obtained when using an extension of her approach to find equivalent statements for simple implicit negations of "all-statements" (see Section I.4 above). The second was her *Distribute, Separate and Substitute* (DSS-) approach extended to the case of simple implicit negations of "there-exist-statements" (see Table 19).

[177] For details on the development of Andrea's assumptions about disproving UASs, see Section I.2.1.2 and Section I.2.2.2 above.
[178] For details on this process, see Section III below.

Table 19. Andrea's conflict when reasoning the simple implicit negation of a negative "there-exist-statement"

Statement	**Approach used**	
	SS-approach	**DSS-approach**
There does not exist X that is not Y	*≡ [There does not exist X] that is not Y*	*≡ ~ (There exists X that is not Y)*[179] *≡ ~ (There exists X) that is ~ (not Y)* *≡ [There does not exist X] that is (Y)*
CONFLICT	*≡ No X is not Y*	*≡ No X is Y*

The following statements played an important role in the excerpt that reveals the way Andrea experienced the conflict.

St123: *There do not exist divisions of natural numbers that have a remainder equal to 0*

St126: *There do not exist divisions of natural numbers that do not have a remainder equal to 0*

St127: *No divisions of natural numbers do not have a remainder equal to 0*

St89: *Not all divisions of natural numbers have a remainder equal to 0*

St125: *No division has a remainder zero*

The excerpt began with Andrea's manifestation of a doubt she had at that moment and exhibits the use of her DSS-approach to negate St126.

> Andrea: *Right now, my doubt is, for example, when I have, hmm, not all, where is it? Above [she means St89], when I have not all, that "not" changes "all" and changes the middle [she means its consequent] ... However, in statement #8 [here St126], when it says "there does not exist", then that "there does not exist" is-... I took that "there does not exist" as "none gives zero" [St125], but I thought, the middle should change as well; however, it does not change. It only changed "there does not exist" to "no".*
>
> I: *It changed to "all [divisions] have remainder 0".*
>
> Andrea: *Sure, but, I mean, for example, above, for example, I thought, where is it?... when I say "there does not exist" is "no", for me, okay?... I translated it like that. Then I thought, "there does not exist", "no", and up to there, "none gives zero" [St125]; whereas up above in the first statement [St89] is, "not all", "some", and over there the middle changed to "do not give zero" ["do not have remainder zero"]; however, below [in St126] it [the consequent] does not change, only the beginning was changed. I mean, that is my confusion.*

Andrea's conflict can be divided into three crucial events: First, because of her DSS-approach, she determined that *"Not all X are Y"* was equivalent to *"Some X are not Y"* as the negator changed both, the quantifier "all" and the consequent[180]. Second, in the statement *"There does not exist X that is not Y"* the negator should have similarly changed both, the quantifier "there exist" and the consequent. According to Andrea, *"There does not exist X that is not Y"* should have been equivalent to *"No X is Y"*, since *"there does not exist"* was the same as *"no"* and the consequent changed from *"not Y"* to *"Y"* (her expectation after using her DSS-approach). Third, the conflict arose as she contrasted

[179] I use the sign ~ to represent negation.
[180] For details, see Section I.4 above.

"No X is not Y", which she had obtained with her SS-approach (St127 in this context), and *"No X is Y"*, which she obtained with her DSS-rule.

Andrea's reasoning and the resulting conflict are summarized in Table 19, where the contrast between the two statements she obtained when using the two different approaches is clear.

Andrea's DSS-approach was her version of the rule I used to explicitly introduce the negation of ESs: *"if I have the negation of an existential statement, it [the negation] is going to affect also both, my antecedent and consequent"* (turn 1 in Episode 17 above). In simple terms, the rule entailed the negation of both the antecedent and the consequent of the statement.

Andrea's explicit reference to her DSS-approach for the case of simple implicit negations of "all-statements" in the excerpt reveals her expectation that the rule also worked in the case of ESs. The main issue with Andrea's DSS-approach was that it cut the statement into parts that were then negated according to her own criteria. This lack of understanding of negation stemmed from my input for the negation of statements. Notably, I suggested to the teachers that, when negating an ES, the negator affected both parts of the statements, without specifying how.

In the following sections I move towards pointing out some aspects that might explain the reasoning behind Andrea's assumptions related to the simple implicit negation of "there-exist-statements". I also include a comparison between "there-exist-statement" and "some-statements" and how this might have influenced Andrea's reasoning about existential statements.

A possible explanation why Andrea might not have found it difficult to disprove a simple implicit negation of an affirmative "there-exist-statement"

Here I point out two possible reasons that might have played an important role in this context: the form of the statement and Andrea's background knowledge.

The form of the statements

Andrea's reasoning about the implicit negation of affirmative "there-exist-statements" might be explained in terms of the form of the statements. Table 20 includes three equivalent existential statements (column 1), their respective implicit negations (column 2), and their equivalent negations (columns 3 and 4).

Table 20. Equivalent Existential Statements, their respective implicit negations and some equivalent negations

<table>
<tr><th>Statement</th><th>Implicit Negations</th><th colspan="2">Equivalent Negations</th></tr>
<tr><td rowspan="2">1. There exists X that is Y</td><td>1.1. There does not exist X that is Y</td><td rowspan="5">No X is Y</td><td rowspan="5">All X are not Y</td></tr>
<tr><td>1.2. It is not the case that there exists X that is Y</td></tr>
<tr><td rowspan="2">2. There is X that is Y</td><td>2.1. There is no X that is Y</td></tr>
<tr><td>2.2. It is not the case that there is X that is Y</td></tr>
<tr><td>3. Some X are Y</td><td>3.1. It is not the case that some X are Y</td></tr>
</table>

In Spanish the form of the statements in Table 20 looks even simpler (see Table 21 below).

Table 21. Spanish equivalents of the statements in Table 20

Proposición	**Negación Implícita**	**Negaciones Equivalentes**	
1. Existen X que son Y	1.1. *No* existe X que sea Y	Ningún X es Y	Todo X no es Y
	1.2. *No es el caso que* existen X que son Y		
2. Hay X que son Y	2.1. *No* hay X que sea Y		
	2.2. *No es el caso que* haya X que sea Y		
3. Algunos X son Y	3.1. *No es el caso que* algunos X son Y		

First, in order to claim the existence of *X* that is *Y*, we say *“Existen X que son Y”* (or *“Hay X que son Y”*) instead of *“There exists X that is Y”* (or *“There is X that is Y”*), which avoids the use of the term “there” (compare statements 1 and 2 in Table 20 and Table 21, respectively, or see Table 22).

Table 22. Comparing “there-exist-statements” in English and Spanish

English	***Spanish***
There exists X that is Y	*Existen X que son Y*
There is X that is Y	*Hay X que son Y*

Second, when formulating an implicit negation of the existence of *X* that is *Y* in Spanish we do not add anything but the negator “no” in front of the statement; that is, we say *“No existe X que sea Y”* (or *“No hay X que sea Y”*) unlike its counterpart in English that involves adding two particles, the negator “not” and the auxiliary “does”. Thus, the implicit negation of *“There exist X that is Y”* becomes *“There does not exist X that is Y”* (compare statements 1.1 and 1.2 in Table 20 and Table 21, respectively; or see Table 23). Hence, in order to prove that *“no existe X…”* (*“There does not exist X…”*) is false in Spanish, it is sufficient to show that *“[sí[181]] existe X…”* (*“There [does] exist X…”*).

Table 23. Comparing implicit negations for an affirmative “there-exist-statement” in English and Spanish

in English	***in Spanish***
There does not exist that is Y	*No existe X que sea Y*
There is no X that is Y	*No hay X que sea Y*

On the other hand, in both languages (English and Spanish) the first and second existential statements admit two implicit negations, unlike the third existential statement (“some-statements”) that only admits one[182]. Statements 1.1 and 2.1 are alternative implicit negations for the first and second statements, respectively, that the third statement does not admit. These statements are simpler negations than 1.2 and 2.2, respectively. The implicit negations 1.1 and 2.1 consist of basically adding the one-word negator “not” to the front of the statement; whereas the implicit negations 1.2 and 2.2 require adding more than one word to the front of the statement in order to negate it. All this means that there are several advantages of having implicit negations of “there-is-” and “there-exist-” statements over “some-statements”. Hence, the simpler form of the two implicit negations (what I called “simple implicit negation”) seems to have supported Andrea’s straightforward justification.

[181] “Sí” in Spanish is “yes” in English. In this context I used it to emphasize the contrast with the negator “no”. It could be interpreted as “yes, indeed, there exists X…”

[182] “Implicit” negations of the form *“It is not the case that…”*, could be also rephrased as *“it is not true that…”* or *“it is false that…”*. I consider only one of those as representative of the others since they have a very similar form.

It is not surprising that Andrea used the same approach to disprove the simple implicit negation *"There is no X that is Y"* by proving that *"There is X that is Y"*.

Andrea's background knowledge

Concluding that an existential statement of the form *"There does not exist X that is Y"* is false would have not been possible by only considering the form of the statement. Certainty about the existence of an *X* that was *Y* (*"There exist X that is Y"* is true) was also a necessity. This, in Andrea's case, was guaranteed by two factors: first, her current mathematical interpretation of the "there-exist-statement", and second, her mathematical background knowledge[183] that ultimately confirmed the existence of an *X* that was *Y*.

These two factors seem to have been the supportive context that favored Andrea's aligned with mathematics *disproving* of the simple implicit negation of *affirmative* "there-exist-statements".

A possible explanation for Andrea's ease to find equivalent statements for the simple implicit negation of "there-exist-statements"

Andrea's "separate and substitute" approach

In Table 24 I include the first equivalent statements that Andrea provided for the simple implicit negations of "there-exist-statements".

Table 24. Simple implicit negations of "there-exist-statements" and their equivalent statements

Simple implicit negation of "there-exist-statements"	Equivalent Statement
There does not exist X that is Y	*No X is Y*
There does not exist X that is not Y	*No X is not Y*

The form of the simple implicit negation of "there-exist-statements" plus her available *semantic substitution* of "there does not exist" with "no/none" (her SS-approach) led Andrea to produce equivalent statements that were mathematically aligned.

Even though the application of this approach yields to equivalent statements, it involves "cutting" statements in non-mathematical ways. While the approach "works well" with both affirmative and negative "there-exist-statements" (as shown in Table 24), it does not with other statements. For example, think of the simple implicit negation of a UAS, say *"Not all X are Y"*. If we use Andrea's SS-approach to find an equivalent statement for it, substituting "not all" with "some", we obtain the statement *"Some X are Y"*[184]. However, this "some-statement" is not mathematically equivalent to *"Not all X are Y"*.

A possible explanation for Andrea's initial assumption dAA9[5] to disprove the simple implicit negation of a negative "there-exist-statement"

Andrea initially accepted that to disprove *"There does not exist X that is not Y"*, an *X* that is *Y* needs to be shown (her assumption dAA9[5]). Her assumption might be explained by her overlooking of the negative consequent in the statement, as if she was evaluating

[183] The background knowledge used to prove the statements involved in the intervention was intended to be mostly built during the first part of the intervention that focused on the mathematical content (see Chapter 4, Section II.2.1, for details on the design of the intervention).
[184] For details, see Section I.4 above.

the statement *"There does not exist X that is Y"* instead (compare her assumptions dAA9[5] and dAA9[2]). The way some double negations are interpreted as simple negations in Spanish can be accountable for Andrea's possible omission.

Double negations in Spanish

Table 25 includes three examples of sentences in Spanish where double negations are employed (first column), the way they are commonly interpreted in Spanish (second column), literal English translations of those statement where the use of the two negation particles can be seen (third column), and finally a "closer" translation into English (fourth column).

Table 25. Examples of double negations in Spanish (Real Academia Española & Asociación de Academias de la Lengua Española [2010, p. 362, 370, 925]).

<table>
<tr><th>Sentence in Spanish</th><th>Its interpretation in Spanish</th><th>A literal translation to English</th><th>A "closer" English translation</th></tr>
<tr><td rowspan="2">1. Ella no tiene ningún interés</td><td rowspan="2">She is not interested at all</td><td>She not have no interest</td><td rowspan="2">She has no interest</td></tr>
<tr><td>She does not have no interest</td></tr>
<tr><td rowspan="2">2. Aquí no hay nada</td><td rowspan="2">A place where somebody just entered is empty</td><td>Here not there is nothing</td><td rowspan="2">Here there is nothing / There is nothing in here</td></tr>
<tr><td>Here there is not nothing</td></tr>
<tr><td rowspan="2">3. No vino nadie</td><td rowspan="2">Nobody came</td><td>Not came no one</td><td rowspan="2">No one came / Nobody came</td></tr>
<tr><td>It did not come nobody</td></tr>
</table>

The first sentence *"Ella **no** tiene **ningún** interés"* contains two negation particles: "no" ("not" in English) and "ningún" ("none" or "no" in English). In Spanish this is usually interpreted as *"Ella **no** tiene interés"* ("the interest she has is zero" or simply "she is not interested at all" in English). The second sentence *"Aquí **no** hay **nada**"* includes the two negation particles "no" ("not" in English) and "nada" ("nothing" in English). It can be interpreted as *"aquí hay **nada**"* ("there is nothing in here" or a "place where somebody just entered is empty" in English). The third sentence *"**No** vino **nadie**"* has the two negation particles "no" ("not" in English) and "nadie" ("nobody" in English). It is usually interpreted as *"**Nadie** vino"* ("nobody came" in English). emphasizes that *not a single person came*.

This phenomenon in Spanish is known as *reinforcement of the negation* (see e.g., Martí, Taulé, Nofre, Marsó, Martín-Valdivia & Jiménez-Zafra, 2016) and consists of the use of a second negation particle in order to reinforce a negation. Even though there are two negation particles included in the statements, they are interpreted as one. This phenomenon might explain why Andrea presumably ignored or omitted the second negation particle when disproving the statement *"There does not exist X that is not Y"*. She might have interpreted the "double negation" as a reinforcement of the negation, and it led her to focus on the statement *"There does not exist X that is Y"* instead.

The treatment of the implicit negation of a negative "there-is-statement" after the intervention suggested that the teachers' were interpreting the statement as a reinforcement of the negation. Specifically, during our Meeting #6 the teachers were asked to find out the truth value of St134 and justify it.

St134: *There are no numbers that are not bigger than 6*

Lizbeth first claimed that St134 could be stated as *"No numbers are bigger than 6"*. Andrea did not show any signs of disagreement, which suggests she agreed with Lizbeth's interpretation.

Whether Andrea in fact assumed that statements of the form *"There are no X that are not Y"* and *"There are No X that are Y"* (or *"There does not exist X that is not Y"* and *"There does not exist X that is Y"*) are equivalent, because of the double-negations phenomenon, is not completely clear to me. Recall that Andrea expressed concern when she obtained two conflicting negations when using two approaches to find equivalent statements for simple implicit negations (see Table 19 in "The conflict" above). She did not seem comfortable with the idea of obtaining *"No X is not Y"* and *"No X is Y"* as two equivalences of the simple implicit negation *"There does not exist X that is not Y"*. If Andrea had used the reinforcement of a negation to interpret the two "no-statements", she would not have been conflicted during the discussion. Nevertheless, other factors might have played a role in Andrea's interpretation.

Even though it is not completely clear whether Andrea was or was not using a reinforcement of the negation, it is clear that Lizbeth was, and as such the interpretation of double negations as one negation deserves attention[185].

A possible explanation for the different performances when negating ESs

The way the tasks were formulated and therefore the kind of attention they drew might have influenced the teachers' performance when solving them.

Tasks with different demands

While Andrea's initial assumptions allowed her to find an expected equivalent statement for *"There does not exist X that is Y"*, she struggled to negate a statement of the form *"Some X are Y"* (see Sections II.3.2.1 and II.3.1 above for details). Observe that both statements, *"There exist X that is Y"* and *"Some X are Y"*, are equivalent existential statements. Hence, finding a negation for *"Some X are Y"* and finding an equivalent statement for *"There does not exist X that is Y"* should generate equivalent statements. Nevertheless, the tasks involved different demands that entailed different challenges. Whereas one task explicitly asked the teachers to *negate* a statement, the other asked them to *find an equivalent statement* for an implicit negation. Furthermore, as I explained before, "some-statements", unlike "there-exist-statements", do not have a simple implicit negation (see Table 20). On one hand, Andrea was not sure what the negation of a "some-statement" was (see Section II.3.1). On the other hand, her "separate and substitute" approach facilitated finding an equivalent statement for the simple implicit negation of an affirmative "there-exist-statement" (her assumption dAA9[4]).

During the intervention I did not include a task that explicitly requested the teachers to negate a "there-exist-statement", as I did for "some-statements". In addition, I did not ask the teachers to find an equivalent statement for implicit negations of "some-statements" like *"It is not the case that some X are Y"*. Further research might be needed to find out whether different requests in tasks play a major role in individuals' performances.

[185] For details on this recommendation, see Chapter 7, Section II.9.9.

Summary of Section II.3

Andrea began the intervention with several initial assumptions related to the negation of existential statements. Most of them referred to the simple implicit negation of "there-exist-statements". Notably, Andrea assumed that a statement of the form *"There does not exist X that is (not) Y"* was equivalent to *"No X is (not) Y"*, which resulted from her semantic substitution of "there does not exist" with "no/none".

The development of Andrea's assumptions for the negation of "some-statements" experienced some challenges. Her struggles were directly related to two possible factors: First, Andrea's initial non-mathematical interpretation of the quantifier "no" and "no-statements". This impeded her from understanding that the negation of a "some-statement" was a universal statement, no matter that it was expressed as an "all-" or "no-" statement. Second, unlike the semantic substitution Andrea used for simple implicit negations of "all-statements" ("not all" means "some"), she did not have a semantic substitution for "not some". In fact, there is no way to express the negation of a "some-statement" as a simple implicit negation.

Unlike "some-statements", "there-exist-statements" admit simple implicit negations (e.g., *"There does not exist X that is [not] Y"*). In particular, Andrea disproved the simple implicit negation of an *affirmative* "there-exist-statement" (*"There does not exist X that is Y"*) and found an equivalent statement for it (*"No X is Y"*) with no difficulties. To disprove the statement, she relied on its simple form (simpler in Spanish), where she needed to focus on showing that there did exist an *X* that was *Y*. To find an equivalent statement, she relied on her *separate and substitute* approach (she substituted "there does not exist" with "no") to obtain *"No X is Y"*.

Andrea's main challenge was revealed as she engaged in similar activities for the case of the simple implicit negation of a *negative* "there-exist-statement". Andrea disproved the statement *"There does not exist X that is not Y"* as if she was disproving the statement *"There does not exist X that is Y"*. This suggested that she might have used a *reinforcement of the negation* to interpret both statements as the same. As a response to my rejection, she used her *separate and substitute* approach to obtain *"No X is not Y"* and *"All X are Y"* as equivalent statements, which she disproved based on her understanding of disproving USs. At that point Andrea faced a conflict given that she obtained two different "no-statements" when negating *"There does not exist X that is not Y"* with the two approaches she seemed to have assumed were correct. When she used her *separate and substitute* approach, she obtained *"No X is not Y"*; whereas when she used her DSS-approach), she obtained *"No X is Y"*. This suggests that Andrea might not have used the reinforcement of a negation to interpret statements (as Lizbeth did) and exhibits her lack of understanding of the negation of existential statements.

Andrea's approaches to find equivalent statements for simple implicit negations of "there-exist-statements" (her *separate and substitute* approach and her *DSS-approach*) were uncertain shortcuts that did not support a mathematical understanding of negations. Both of them were strongly rooted in the semantic substitution of "there does not exist" with "no", which involved cutting the statement in non-mathematical ways.

Figure 45 shows the development of Andrea's assumptions related to the negation of ESs and the elements or aspects of the intervention that might have supported or triggered changes[186].

[186] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

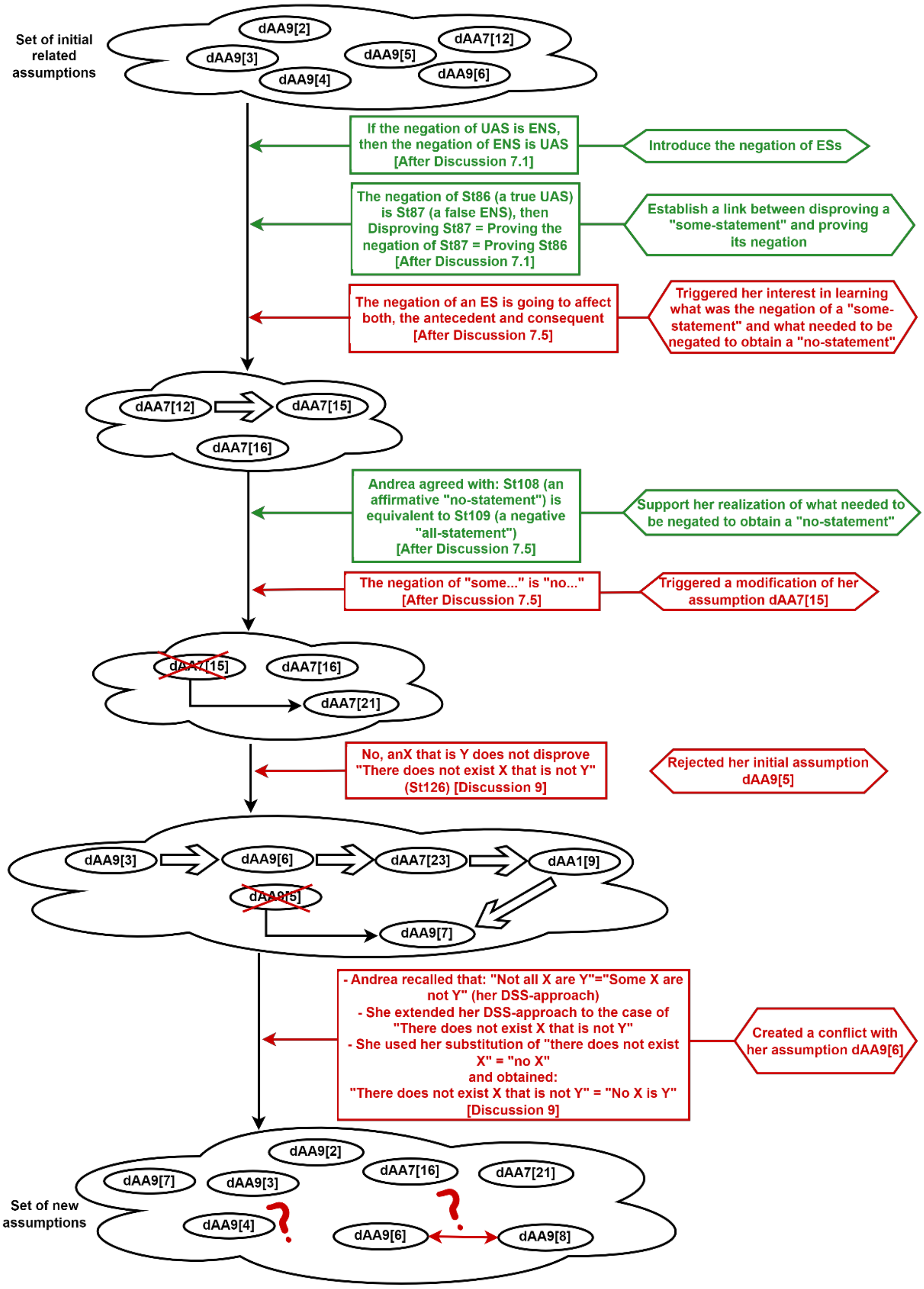


Figure 45. The development of Andrea's assumptions about the negation of ESs (dAA9[2]: "There does not exist X that is Y" is false because there is indeed an X that is Y (i.e., because "There exist X that is Y" is true); dAA9[3]: There does not exist X ≡ No X; dAA9[4]: "There does not exist X that is Y" is the same as "No X is Y"; dAA9[5]: "There does not exist X that is not Y" is false because there exist an X that is Y; dAA9[6]: "There does not exist X that is not Y" is the same as "No X is not Y"; dAA7[15]: The negation of a "some-statement" is an "all-statement", and never a "no-statement"; dAA7[16]: The negation of "Some X are Y" ("Some X are not Y") is "All X are not Y" ("All X are Y"); dAA7[21]: The negation of "Some X are Y" is "No X is Y"; dAA9[7]: "There does not exist X that is not Y" is false because there is an X that is not Y; dAA9[8]: "There does not exist X that is not Y" is the same as "No X is Y"; dAA7[12]: "No X is Y" and "No X is not Y" refers to everything that is not X; dAA7[23]: "No X is not Y" is equivalent to "All X are Y"; dAA1[9]: A counterexample must satisfy the first condition, but contradict the second condition of the universal statement)

III. The teachers' assumptions about Universal Negative Statements

In this section I focus only on the development of Andrea's assumptions related to universal negative statements (UNSs) as again she participated more than the others, and explicitly displayed her thinking process.

Universal Negative Statements can be presented in different forms; however, the most common form is *"No X is Y"* (an affirmative "no-statement"), which is equivalent to *"All X are not Y"* (a negative "all-statement"). In Table 26 I provide a comparison of affirmative "no-statements" and a negative "all-statement" in English and Spanish.

Table 26. Equivalent Universal Negative Statements in English and Spanish

Affirmative "no-statements"		***Negative "all-statement"***	
in English	***in Spanish***	***in English***	***in Spanish***
No X is Y *None of X is Y*	*Ningún X es Y* *Ninguno de los X es Y*	*All X are* ***not*** *Y*	*Todos los X* ***no*** *son Y*

I separate this section into two main sub-sections: "no-statements" (Section 1), and negative "all-statements" (Section 2), as Andrea made a clear distinction between these kinds of statements.

Section 1 is divided into three main topics: interpretation of "no-statements" (Section 1.1); disproving of "no-statements" (Section 1.2); and negation of "no-statements" (Section 1.3). Likewise, Section 2 has a focus on three similar themes in relation to negative "all-statements": their interpretation (Section 2.1); their disproving (Section 2.2); and their negation (Section 2.3).

1. "No-statements"

Here I use the quantifiers "no" and "none" interchangeably. Hence, I do not distinguish "no-statements" and "none-statements". I call a statement of the form *"No X is Y"*, an *affirmative* "no-statement" and a statement of the form *"No X is not Y"*, a *negative* "no-statement".

The core discussions about "no-statements" took place right after Discussion 7.5. They were motivated by Andrea's interest in learning what the negations of "some-statements" were like. At that point Andrea had already used at least two of her initial assumptions related to "some-statements" (1) *"Some X are Y"* necessarily implies that *"Some X are not Y"* (dAA7[3]), and (2) *"Some X are Y"* and *"All X are Y"* cannot be both true (dAA1[1])[187].

1.1. Interpretation of "no-statements"

Andrea's assumptions related to "no-statements" and their interpretation involved her use of a non-mathematical meaning of the quantifier "no". Her initial assumptions were exhibited not only through her interpretation of "no-statements", but also through her representation of them.

Table 27 includes Andrea's initial assumptions related to the interpretation of "no-statements" as they were first observed during and after the intervention.

[187] For details, see Section II.1 above.

Table 27. Andrea's assumptions related to the interpretation of "no-statements"

Assumption	**Type of assumption (when first observed)**
dAA7[12]: *"No X is Y"* and *"No X is not Y"* refers to everything that is not *X*	Initial assumption (during the intervention)
dAA7[13]: *"No X is Y"* is represented as two disjoint sets *X* and *Y*	Initial assumption (during the intervention)
dAA7[20]: *"No X is Y"* is different from *"All X are not Y"*	Initial assumption (during the intervention)
dAA7[25]: *"No X is not Y"* is represented as two disjoint sets *X* and *Y*	Initial assumption (during the intervention)
aAAm3: Given two disjoint sets *X* and *Y*, *Y* represents the set of the elements that are "not X"	Initial assumption (after the intervention)
dAA7[14]: *"No X is Y"* is equivalent to *"All X are not Y"*	New assumption (during the intervention)
dAA7[23]: *"No X is not Y"* is equivalent to *"All X are Y"*	New assumption (during the intervention)

Even though her four initial assumptions are closely related, I separate them because it is not clear whether Andrea saw the connections between them.

In terms of her two new assumptions, Andrea used these assumptions in parallel with her initial assumptions. This suggested that Andrea did not really abandon any of her initial assumptions, as I show below.

I organize this section into two main parts. In the first part I focus on Andrea's initial assumptions related to "no-statements", their interpretation and representation. All these assumptions are related to both affirmative and negative "no-statements". In the second part I focus on Andrea's constant lack of association between "no-" and "all-" statements. I illustrate different contexts where this lack of association was evident and provide possible explanations for it.

Andrea's initial assumptions

In this first part I include Andrea's four initial assumptions. The three first assumptions were identified during the intervention for teachers, while the fourth was identified after that, during one of our meetings. Assumption dAA7[12] involves Andrea's initial interpretation of "no X…". Assumptions dAA7[13] and dAA7[25] are about Andrea's initial representation for affirmative and negative "no-statements". Assumption aAAm3 is related to Andrea's initial interpretation of the simple negation "not X".

"No X is Y" and "No X is not Y" refers to everything that is not X (dAA7[12])

Andrea's initial assumption dAA7[12] involved the elements Andrea presupposed that "no-statements" referred to. Andrea asserted that *"no X..."* was about the elements outside set *X*; that is, the elements that were not in *X*.

Motivated by Andrea's interest in "no-statements", Discussion 7.5 was immediately followed by the analysis of three "no-statements" (St108, St115 and St117, see below). The analyses of these three examples are illustrations of Andrea's use of her initial assumption dAA7[12]. I asked the teachers whether St108 referred to *at least one* odd number, *all* odd numbers, or what exactly the statement referred to.

St108: *No odd number is an Innova number*

Andrea explicitly claimed that St108 referred to *"none… not even one odd number… zero [of them]"*. With a gesture of a fist to represent the set of odd numbers, Andrea used her other hand to point to what was outside her fist, as if she were "sweeping" with that hand (her "fist & sweeping" gesture), to represent the elements outside the set of odd numbers, which according to her, referred to "no odd numbers" (see Figure 46).

Figure 46. Andrea's use of her "fist & sweeping" gesture to explain what she means by "no odd number…"

Andrea used the same assumption for her interpretation of the two negative "no-statements" St115 and St117.

St115: *No odd number is not an Innova number*

St117: *No man is not a living being*

For the case of St117 Andrea explained the following:

> Andrea: *No man, I draw my group of men, none of them, that means those outside, they are not living beings… if I say no man, it means not these [the elements in the set of men] (she uses her "fist & sweeping" gesture again), but those outside… those that are not men… those outside are not living beings.*

Observe that unlike statements St108 and St115, which were imaginary statements, St117 was a true familiar-context statement. Andrea's use of her assumption dAA7[12] for St117 shows that dAA7[12] was a stable assumption that did not depend on the context and the truth value of the "no-statement".

Andrea's initial interpretation of "no-statements" and elements they refer to differs from a mathematical interpretation. Moreover, Andrea's interpretation is not aligned with a common-language interpretation. From both mathematics and common language perspectives, "no-statements" of the form *"No X…"* are about all the elements in *X*. For instance, the mathematical "no-statement" *"No even number is palindrome"* is about *all even numbers* and it means that every even number is not a palindrome. Likewise, in an ordinary conversation context, the expression "no one came" is interpreted as *every one did not come* (Abrusci et al., 2016, p. 188).

A possible explanation for Andrea's initial assumption dAA7[12]

There are two aspects that might explain Andrea's initial interpretation of affirmative and negative "no-statements": (1) the meaning of "none/no" and (2) the meaning of "not", both in an everyday-life context.

Some of the meanings of the quantifier "no/none" in Spanish are related to Andrea's initial interpretation of "no X…". In concrete, "ningún/ninguno(a)" ("no/none" in

English) can express the non-existence of that denoted by the name it modifies[188]. For example,

- "No he tenido *ningún* problema" ("I had no problems" in English) is interpreted as *I had zero problems*;
- "Ayer compré chocolates, pero no queda *ninguno*" ("Yesterday I bought chocolates, but there are none left" in English) is construed as *there are zero chocolates left*;
- "*Ninguna* de tus amigas estaba de acuerdo" ("None of your friends agreed" in English) is interpreted as *zero of your friends agreed.*

In these cases, the meaning of "none/no" is linked to Andrea's interpretation of "no X…" as "zero elements in X". It is possible that Andrea used her everyday-language meaning of the quantifier "none/no", for example, to identify that the statement *"No odd number is an Innova number"* (St108) refers to *"none… not even one odd number… zero [of them]"*.

In addition, in Spanish there are also some similarities between the quantifier "ningún/ninguno(a)" ("none/no" in English) and the negator "no" ("not" in English) that might explain Andrea's initial interpretation of "no X…" as "the elements outside X, those that are not X". For example, in the context of a girl talking to her mother about her birthday party plans, the girl asks:

"¿A quién puedo invitar a mi fiesta?" ("Who can I invite to my party?" in English),

and the mother replies,

"A quien quieras, pero *no* niños" ("Whoever you want, but *not* boys" in English).

The mother's answer means that the girl can invite others than boys, but no boys, zero boys, which suggests "only girls". In this sense, it is possible that Andrea used this interpretation to conclude that the statement *"No man is not a living being"* (St117) refers to *"those that are not men"* (see above).

"No X is Y" can be represented as disjoint sets X and Y (dAA7[13])

Another initial consideration Andrea used about "no-statements" was related to their representation. Andrea assumed that the diagram that represented the statement *"No X is Y"* consisted of two disjoint sets: *X* and *Y* (her initial assumption dAA7[13]). Her assumption was first observed after Discussion 7.5, during the intervention, when the first "no-statement" St108: *"No odd number is an Innova number"* was put forward. Specifically, I asked the teachers what diagram represented statement St108. Andrea claimed that the (Euler) diagram consisted of two "separated sets", the set of odd numbers (O) and the set of Innova numbers (I); in other words, the sets had an empty intersection. I drew the diagram in Figure 47 and Andrea agreed that it represented statement St108.

[188] Diccionario de la Real Academia Española. Link: https://dle.rae.es/ninguno?m=form

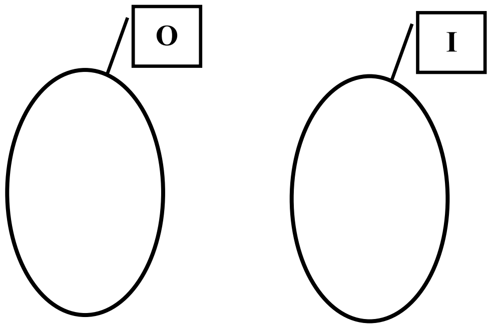


Figure 47. A recreation of the diagram I drew on the whiteboard based on Andrea's directions for the representation of "No odd number is an Innova number" (O denotes the set of odd numbers and I represents the set of Innova numbers)

"No X is not Y" can be represented as disjoint sets X and Y (dAA7[25])

Andrea introduced statement St117 to discuss its representation.

> St117: *No man is not a living being*

The discussion exhibited Andrea's initial assumption (dAA7[13]), for how to represent a negative "no-statement". Andrea claimed that the representation for St117 consisted of the disjoint sets M (set of men) and LB (set of living beings) (see Figure 48).

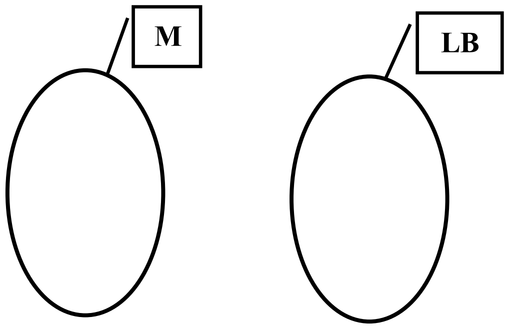


Figure 48. Andrea's expected representation for "No man is not a living being" (St117)

She explained her rationale as follows:

> Andrea[189]: *But if I say no man, it means not these [elements in the set of men], but those outside (she uses her "fist & sweeping" gesture again[190]) … those that are not men… If I am told no man, those are the ones outside. That means that the ones outside are not living beings. Therefore, the group of living beings cannot be next to the one [set] of men… that is why they [the set of men and the set of living beings] should be separated.*

With her claim that *"the group of living beings cannot be next to the one [set] of men"* Andrea suggested that sets M and LB should not be overlapping.

[189] This is a more complete version of the excerpt that I included in Andrea's initial assumption dAA7[12] (above). Further details can be found in the section "The conflict" below.
[190] See Figure 46 above.

Furthermore, the diagram I drew on the whiteboard by following Andrea's directions suggests that Andrea's attention was placed on the elements outside M and outside LB; that is, Andrea's focus was on the blue background as shown in Figure 49. This makes sense according to her initial interpretation of "no X…". For Andrea, "no man" referred to the elements outside M and those elements were not living beings, so they should be also outside LB.

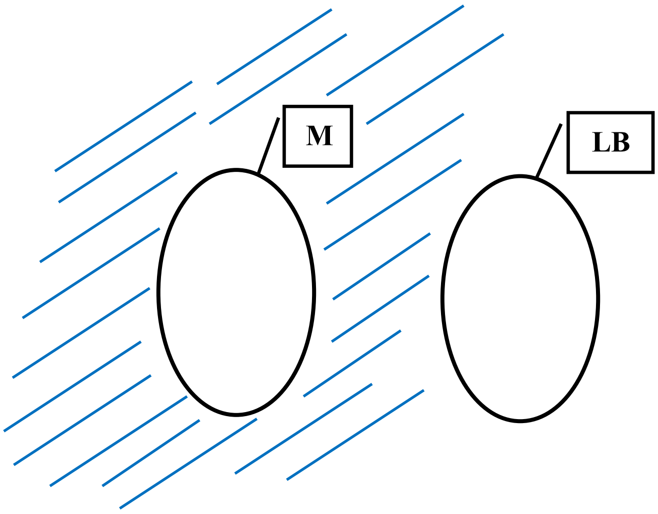


Figure 49. A reconstruction of the diagram I drew on the whiteboard for "No man is not a living being" under Andrea's directions.

This means that for both affirmative and negative "no-statements" Andrea's representation was the same: two disjoint sets. However, because of her initial assumption dAA7[12], her attention in each case was placed on different parts of the diagram (as shown in Figure 50). This is confirmed much later with Andrea's initial interpretation of "not X" (see her assumption aAAm3 below).

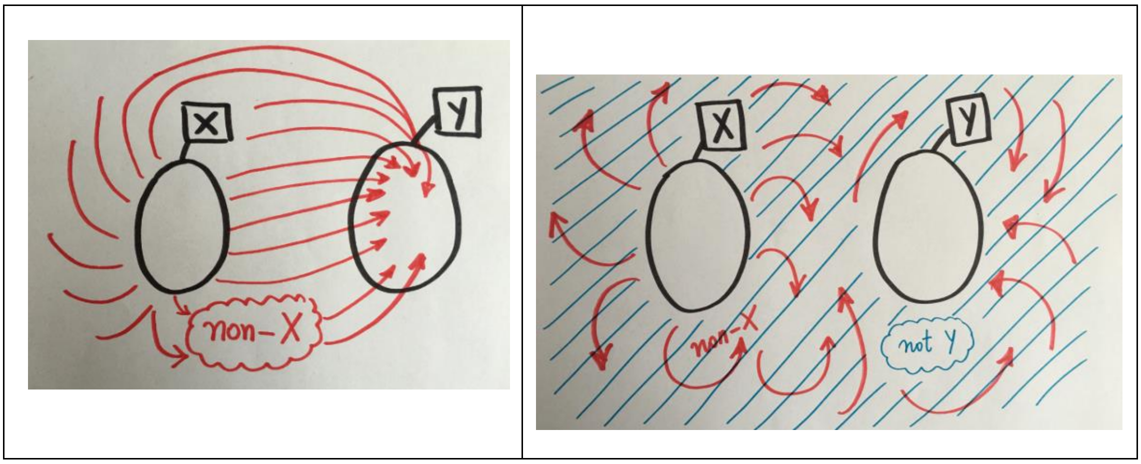


Figure 50. Andrea's attention when representing the affirmative "no-statement" "No X is Y" (on the left) and the negative "no-statement" "No X is not Y" (on the right)

Observe that for Andrea *"No X…"* referred to the elements outside X or those that were *"not X"*. In that sense, she first focused on the elements that were outside X and then she focused on whether those elements should belong to the consequent set or not, according to whether the consequent was affirmative or negative. For example, when Andrea represented *"No X is Y"* she first focused on the elements that were not *X* and then that they *should be* in Y, because the consequent was affirmative (the elements outside X *are* Y). In the same vein, when she represented *"No X is not Y"* she first focused on the elements that were not X and then that they *should not be* in Y, because the consequent was negative (the elements outside X *are not* Y). In other words, she interpreted the statement *"No X is (not) Y"* as if it indicated the set *"Not X and (not) Y"*. This would mean that in the case of her representation for St108 (*"No odd number is an Innova number"*, see Figure 47), Andrea's attention was on the elements in the set of Innova numbers (set I) and not in the set of odd numbers.

Andrea's form of reasoning the representation about "no-statements" shows that despite the fact that Andrea provided the expected representation for *"No X is Y"* (disjoint sets *X* and *Y*), her rationale was grounded on a non-mathematical assumption (Andrea's assumption dAA7[12]).

Given two disjoint sets X and Y, Y represents the set of the elements that are "not X" (aAAm3)

Andrea's initial assumption aAAm3[191] is about the way Andrea interpreted a simple negation like *"not X"*. Andrea assumed that given two disjoint sets *X* and *Y*, *Y* represented the set of the elements that were *not X*. Andrea's assumption was explicit during Meeting #12, after the intervention. The teachers were asked to solve a task[192] that consisted of two sets *P* and *N,* where *P* was the set of "perfect" numbers and *N* was the set of "nephew" numbers (the task did not specify what "perfect" and "nephew" numbers were). The teachers were given Figure 51 as the diagram that represented the relationship of both sets, *P* and *N*. The teachers were requested to color the region where the numbers that were not "perfect" numbers could be located.

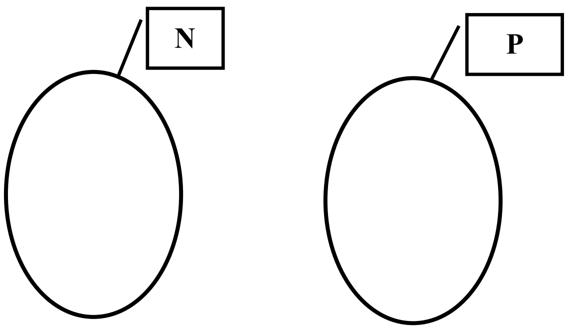


Figure 51. The Euler diagram that was given to show the way sets N and P were related, as shown in Task 5 of the Extra Activity 2.

Andrea colored only set *N* (as in Figure 52), which clearly manifested her current interpretation of *"not P"* in that context, that is, *if a number is not "perfect", then it must be "nephew"*.

[191] I presume that this is an *initial* assumption because Andrea did not take much time to provide an answer and at this point she had not changed yet her assumptions related to "no-statements" and what elements they referred to (her assumption dAA7[12]).

[192] For details, see Task 5 of the Extra Activity 2 (see Appendix EA2).

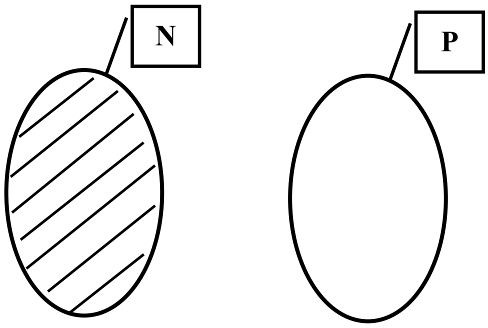


Figure 52. A reconstruction of Andrea's solution to task 5 of the Extra Activity 2

Andrea seems to have overlooked that outside sets N and P there *may* be elements. She might not have colored the region outside N and P because she assumed that it was empty. Actually, when I pointed out that outside N and P there may be elements, she was surprised, but did not object and accepted it.

Andrea's solution of the task fits her representation for *"No perfect number is a nephew number"*. If she had in mind the negative "no-statement" *"No perfect number is not a nephew number"*, she would have colored the elements outside N and P, but she did not.

Permanent lack of association between "no-statements" and "all-statements"

I supported the teachers' understanding of "no-statements" by drawing their attention to the equivalence with "all-statements". I expected the teachers to use what they had learned about "all-statements" and transfer those insights when learning about "no-statements". In other words, I expected "all-statements" to be used as *"bridge statements"* to the understanding of "no-statements". However, Andrea did not manage to see the connection between "no-statements" and "all-statements". She chose to reason directly with "no-statements" and avoid their equivalent "all-statements".

During the intervention there were three contexts where Andrea's lack of association between "no-" and "all-" statements was evident: the discussion of the negation of "some-statements", the discussion of the negation of "no-statements" and the discussion of the representation for negative "no-statements". Here I focus on the first two. I split this section according to the two contexts to illustrate Andrea's lack of association between "no-" and "all-" statements. I include the third context (A conflict: Different representations for *"No X is not Y"*) in its own section as a way to draw attention to a conflict that Andrea faced when representing negative "no-statements".

The negation of "some-statements"

The first piece of evidence emerged as soon as the negation of existential statements was introduced right after Discussion 7.5. The following is an excerpt that exhibits Andrea's lack of awareness of what results when negating a "some-statement" and her lack of association between "no-" and "all-" statements. This is the first discussion about "no-statements" we had during the intervention. It emerged as a result of Andrea's interest in finding out whether an "all-statement" or a "no-statement" was obtained when negating a "some-statement"[193].

[193] For more details, see Section III.3 below.

Andrea: *I mean, if I negate that "some", does it become "all" or "none"?*

I: *If you negate this one (I point to St106: Some natural numbers divisible by 2 are not divisible by 3)? What happens if we negate this one [St106]?*

Andrea: *It is going to become "all".*

I: *How is it going to change? All numbers-*

Gessenia: *Divisible by 2 are divisible by 3.*

Andrea: *Alright, then it will never become "none".*

Andrea's enquiry about whether a result of negating a "some-statement" was an "all-" or "no-" statement suggests she made a distinction between the statements. Furthermore, Andrea assumed that since they obtained an "all-statement" when negating the "some-statement" St106, the negation of "some-statements" will never be a "none-statement". With her assertion *"then it will never become 'none'"* Andrea established a boundary between "all-" and "no-" statements[194].

The negation of "no-statements"

Further evidence of Andrea's lack of association between "all-" and "no-" statements was observed when Andrea rejected that negating *"All X are not Y"* was the same as negating *"No X is Y"*.

To provide some context, the equivalence *"No X is Y"* ≡ *"All X are not Y"* first arose and was accepted when we discussed the "no-statement" St108:

St108: *No odd number is an Innova number*

Andrea claimed that St108 referred to the elements outside the set of odd numbers (her assumption dAA7[12]) and that both sets involved in the statement, the set of odd numbers and the set of Innova numbers, should be separate in its representation (her assumption dAA7[25]). Based on the diagram she suggested, I drew the teachers' attention to the odd numbers and asked them what was stated in St108 about those numbers. Andrea answered that *"they are not Innova numbers"* and agreed that St108 and St109 were equivalent statements.

St109: *All odd numbers are not Innova numbers*

This was the first time Andrea acknowledged the equivalence between a "no-" and an "all-" statement during the intervention for teachers. However, her acceptance was only superficial given that she continued to use her initial assumption that *"no X..." referred to the elements that were not X* throughout the intervention and after it. This superficiality was shown later when the teachers were asked to negate St108. Andrea provided St113:

St113: *Some odd numbers are not Innova*

Given that St113 was not the negation of St108, Andrea was asked to negate St109, which Andrea had accepted as being equivalent to St108. If both statements are equivalent, the negation of St109 should be the same as the negation of St108.

Andrea: *Some odd numbers are Innova numbers. But that [negating St109] is different from negating the other statement [St108].*

[194] For more details, see Episode 17 above in section II.3.1.

Andrea provided the statement *"Some odd numbers are Innova numbers"* as the negation of St109, as expected; however, she anticipated and objected to my request to negate St109 as a way to negate St108. Andrea observed that negating St109 was not the same as negating St108, which also exhibited her lack of association between a "no-statement" and its respective equivalent "all-statement".

Andrea's recurrent unwillingness to rely on the equivalence strongly supports my view that Andrea had only accepted the equivalence at a superficial level. Her lack of confidence in the equivalence between "no-" and "all-" statements was shown, for instance, when Andrea negated the "some-statement" St114 (*Some odd numbers are Innova numbers*), obtained *"All odd numbers are not Innova numbers"* and claimed that the negation was not a "no-statement". Even though the negation was not a "no-statement" as their form of expression was not the same, it was equivalent to the "no-statement" *"No odd number is an Innova number"*. In fact, Andrea's refusal to accept that the negation she obtained was a "no-statement" suggested that she rejected the equivalence between *"All odd numbers are not Innova numbers"* and *"No odd number is an Innova number"*.

A conflict: Different representations for "No X is not Y"

A conflict emerged as we engaged in a discussion about the representation for the negative "no-statement" St115:

St115: *No odd number is not an Innova number.*

The conflict arose as Andrea learned that the expected representation for a statement of the form *"No X is not Y"* consisted of set *X* included in set *Y* and not, as she initially assumed, disjoint sets *X* and *Y* (her initial assumption dAA7[25]).

Andrea spontaneously provided the equivalent "all-statement" for St115; that is: *"All odd numbers are Innova numbers"*. Recall that she had accepted the equivalence between "all-statements" and "no-statements" in a superficial way, without any understanding. The equivalent "all-statement" was used to find the negation for St115[195]. Attention was then drawn to finding a representation for St115. Andrea at first agreed with my initial deceptive suggestion of two disjoint sets (the set of odd numbers and the set of Innova numbers). As I reoriented the teachers' attention towards the already analyzed St108 (*"No odd number is an Innova number"*) and its representation, Andrea decided to change her initial answer. She claimed now, though with clear signs of indecision, that the diagram that represented St115 consisted set O inside set I, which I drew on the whiteboard (see Figure 53).

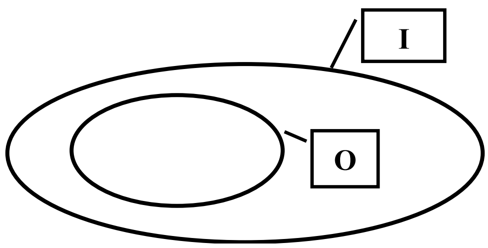


Figure 53. Andrea's suggested representation for "No odd number is not an Innova number" (St115)

Andrea hesitated about the new diagram,

[195] For details see Section III.3 below.

> Andrea: *the set of odd numbers included in the set of Innova numbers. No, no, no. Yes, yes, yes. Uh-huh.*

It is likely that to obtain the new diagram (in Figure 53) Andrea switched her focus towards the equivalent "all-statement" for St115. She represented a diagram for it, instead of one directly for St115. Andrea's vacillation revealed the conflict Andrea experienced at that moment.

Further evidence of her conflict emerged afterwards, when Andrea suggested analyzing another negative "no-statement". Andrea wanted to test her previous conclusion in relation to the representation of St115. To do so she introduced the familiar-context negative "no-statement" St117:

> St117: *No man is not a living being.*

As with St115, Andrea immediately translated St117 into its equivalent "all-statement" St118.

> St118: *All men are living beings*

Then she said:

> Andrea: *How can I represent this statement [St117]? I mean, I forget about this [St118], only with this one [St117]. No man, I draw my group of men, none of them, that means those outside, they are not living beings, then* ***they cannot be inside the (set of) living beings****; however, over there (she points to St118 on the whiteboard) it is, one [set] is inside the other [set], do you understand? I mean, if I focus on all men are living beings, the [set of] men are inside the [set of] living beings (she gestured an inclusion relationship) ... alright, but if I say no man, it means not these [elements in the set of men], but those outside (she uses her "fist & sweeping" gesture again[196]) ... those that are not men... If I am told no man, those are the ones outside. That means that the ones outside are not living beings. Therefore,* ***the group of living beings cannot be next to the one [set] of men****... that is why they [the set of men and the set of living beings] should be separated.*

Andrea clearly experienced a conflict between her expected representation of St117 (two disjoint sets *M* and *LB*) and the representation she would have obtained if she relied on the equivalent "all-statement" St118 (*M* included in *LB*). She first suggested avoiding the equivalent "all-statement" St118 and representing St117 directly (*"I mean, I forget about this [St118], only with this one [St117]"*). Her direct representation for St117 was based on her initial assumption dAA7[12] (*"No man, I draw my group of men, none of them, that means those outside, they are not living beings, then* ***they cannot be inside the (set of) living beings****"*). In fact, she resorted back to her initial interpretation of "no X…" in two additional occasions (*"if I say no man, it means not these [elements in the set of men], but those outside [she uses her "fist & sweeping" gesture again]"* and *"If I am told no man, those are the ones outside. That means that the ones outside are not living beings"*). Andrea knew that the representation for St118 would have *M* included in *LB* (*"if I focus on all men are living beings, the [set of] men are inside the [set of] living beings [she gestured an inclusion relationship]"*); however, she rejected it as the representation for St117 (*"Therefore,* ***the group of living beings cannot be next to the one [set] of men****"*). Andrea "resolved" her conflict by deciding to rely on her own initial assumption dAA7[12], which resulted in two disjoint sets (*"that is why they [the set of men and the set of living beings] should be separated"*).

[196] See Figure 46 above.

Andrea's rejection of the conflicting inclusion representation (M included in LB) in response to the conflict is similar to "rejecting anomalous data" discussed by Chinn and Brewer (1993) and "stonewalling" discussed by Chan, et al. (1997) as ways to respond to cognitive conflicts[197]. Andrea believed that the inclusion representation for St117 was not the correct one. Andrea explained her rejection by relying on her initial assumptions[198]. Notably, she still relied on her fundamental initial assumption that "no X…" refers to the elements that are not X (her assumption dAA7[12]). The use of her initial assumption dAA7[12] during the discussion revealed that she had not modified the main core of her set of initial assumptions.

Table 28 summarizes Andrea's pattern of reasoning to represent both negative "no-statements" St115 and St117. The table includes the conflict Andrea faced and the way she "resolved" it.

Table 28. Andrea's pattern of reasoning for the representation of two negative "no-statements".

Remark #	Andrea's remarks	Andrea's pattern of reasoning
	St115: No odd number is not an Innova number	
1	*"That is equivalent to 'All odd numbers are Innova numbers'"*	Acknowledged the equivalence
2	*"Yes [the representation for St115 is two separated sets]"*	Accepted a disjoint representation for *"No X is not Y"*
3	*"the set of odd numbers included in the set of Innova numbers. No, no, no. Yes, yes, yes. Uh-huh."*	Suggested an inclusion representation for St115 **Conflict:** hesitation about the inclusion representation
4	*"If I say 'no odd number' this means the elements outside the set of odd numbers, right?"*	Resorted to her initial interpretation of "no X…" to evaluate the representation for St115
5	*"What about a real (context) example?"*	Suggested using a familiar example
	St117: No man is not a living being	
6	*"This is equivalent to 'All men are living beings'"*	Acknowledged the equivalence
7	*"I forget about this [its equivalent 'all-statement'], only with this one [St117]"*	Decided to ignore the equivalent "all-statement" and only focus on St117
8	*"No man… those outside, they are not living beings… then they cannot be inside the [set of] living beings"*	Resorted to her initial interpretation of St117 to invalidate the inclusion representation for St117
9	*"however, over there (she points to its equivalent 'all-statement' on the whiteboard) it is, one [set] is inside the other [set], do you understand?"*	**Conflict:** different representations for the equivalent "all-statement" and St117
10	*"I mean, if I focus on all men are living beings, the [set of] men are inside the [set of] living beings"*	Showed awareness of the representation for the equivalent "all-statement"
11	*"alright, but if I say no man, it means not these [elements in the set of men], but those outside ... those that are not men… That means that the ones outside are not living beings. Therefore, the group of living beings cannot include the other [set]…"*	Resorted to her initial interpretation of "no X…" to reject again an inclusion representation for St117
12	*"that is why they should be separated [sets]"*	Suggested a disjoint representation for St117

[197] For details on types of responses when facing a conflict, see Chapter 3, Section I.1.2.

[198] Her initial assumption dAA7[25] for the representation of negative "no-statements" was based on dAA7[12] (see above).

While remarks #1 and #6 show Andrea's acknowledgement of the equivalence between "no-" and "all-" statements, her final resolution of the conflict (remark #12) reveals the limited impact it had on the set of Andreas' initial assumptions. Andrea decided to ignore the equivalence (remark #7) and resorted to her initial interpretation of "no X…" (remarks #8 and #11) to resolve the conflict.

Andrea's weak association of "no-" and "all-" statements is here suggested by her decision to overlook the equivalence and, instead, reason with her initial interpretation of "no-statements" (her assumption dAA7[12]).

A possible explanation for Andrea's lack of association between "no-statements" and "all-statements"

Andrea's rejection of the equivalence between the statements *"No X is Y"* and *"All X are not Y"* can be explained in terms of her initial assumption about what elements these statements referred to. Andrea initially assumed that *"No X is Y"* referred to *none* of the elements in *X*, but the elements outside *X* (her assumption dAA7[12]). On the other hand, as I will show later (in Section III.2.2.1 below), *"All X are not Y"* referred to *some* elements in *X*. The latter stems from the combination of two of Andrea's initial assumption: (1) *"All X are not Y"* and *"Not all X are Y"* are the same, and (2) *"Not all X are Y"* is equivalent to *"Some X..."*[199].

1.2. Disproving "no-statements"

In this section I focus on Andrea's assumption related to disproving *affirmative* "no-statements". Andrea showed evidence of having a clear stance on how to disprove statements of the form *"No X is Y"*, which was in harmony with mathematics.

Table 29 includes Andrea's assumption to disprove affirmative "no-statements", the statements she disproved with her assumption, and whether the statement was disproved during or after the intervention.

Table 29. Andrea's assumption for disproving affirmative "no-statements"

Andrea's assumption for disproving affirmative "no-statements"	**Statements disproved**	**When it was disproved**
"No X is Y" is false because there is an *X* that is *Y*	St125: *No division of natural numbers has a remainder zero*	During the intervention
	St136: *No number is odd*	After the intervention
	St152: *No natural number is divisible by 4*	After the intervention

Andrea's disproving of "no-statements" was observed at three different times. During Discussion 9, the teachers solved Activity 13 that included a discussion of the truth value and respective justification for eight statements[200]. Statement St125 (*"No division of natural numbers has a remainder zero"*) was included among those statements, and was

[199] Recall that Andrea's initial assumption about the equivalence for *"Not all X are Y"* changed during the intervention. Hence, it is important to define when *"Not all X are Y"* is included in order to define whether its equivalent "some-statement" was affirmative or negative. See Section I.4 above for details.
[200] For details on the content of the activity, see Appendix A13.

easily disproved by the teachers. Gessenia provided the *division of 10 by 5* to justify its falsity. Andrea agreed with her and provided a further explanation.

> Andrea: *That [10 divided by 5] is a counterexample as it breaks the relation… because it [10 divided by 5] holds and [St125] says "none", then it is false.*

Andrea used again her assumption after the intervention. During Meeting #6 Andrea suggested the number 7 (an odd number) to disprove St136. Similarly, during her teaching of Session #14 Andrea explained to her class that to disprove St152 an example that "broke" the relation in the statement St152 was required.

> Andrea: *There is something you should have clear in mind… when I have statements that state "there is", what should I show?... but here, it says NONE, … and it is false. Here you need to provide an example that breaks this relation.*

Andrea accepted the number 4 as a justification that disproved St152; however, she included 44 to justify that St152 was false in a clean version of the justification on the whiteboard (see Figure 54).

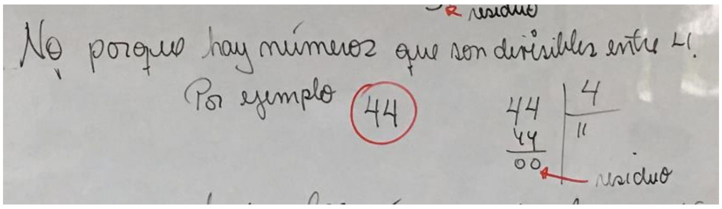


Figure 54. It reads "No, because there are numbers that are divisible by 4. For example: 44".

In all three cases, Andrea was consistent with her assumption about disproving an affirmative "no-statement". Nonetheless, it is not certain whether this assumption was an effect of the intervention (intervention-based assumption) or one of her initial assumptions applied in this context. On one hand, during the intervention I constantly emphasized the equivalence between "no-" and "all-" statements. This might have explained why she disproved affirmative "no-statements" with examples that "broke the relation", as she did when referring to the counterexamples that disproved "all-statements"[201]. On the other hand, during the intervention Andrea usually avoided the equivalence between "no-" and "all-" statements and reasoned with her initial assumptions about "no-statements"[202].

The fact that Andrea took the time to contrast "no-statements" and "there-is-statements" as part of her feedback during her teaching of Session #14 might suggest that she based her reasoning on her initial assumptions. Notably, this episode seems to be tied to Andrea's initial *semantic substitution* of "there does not exist" or "there are no" with "none/no". It is possible that she presupposed that "none…" was false since "there existed…" or "there were…", as she did to disprove simple implicit negations of affirmative "there-exist-statements"[203].

1.3. Negation of "no-statements"

Here I describe the development of Andrea's assumptions related to the negation of "no-statements" and the factors that might have influenced it. The difficulties Andrea faced

[201] For details on the disproving of USs, see Section I.2.1.2 above.
[202] For details on this issue, see Section III.1.1 above.
[203] For details on the disproving of simple implicit negations of "there-is-statements", see Section II.3.2.1 above.

when determining the negation of "no-statements" were first manifested in the discussion for the negation of "some-statements". One important factor in Andrea's struggle to negate "no-statements" stemmed from her lack of a meaningful association between "no-" and "all-" statements. Andrea experienced a conflict when negating "no-statements" and obtaining two different negations.

I split this section in two main parts that focus, first, on the way "no-statements" and their negations emerged in our discussions; that is, the negation of "some-statements", and second, the conflict Andrea faced when negating "no-statements", which revealed interesting assumptions she made about the negation of statements in general.

The negation of "some-statements": Its linkage with "no-statements"

In Section II I included some context on how "no-statements" and their negation emerged in our discussions. Recall that, as we began to discuss the negation of existential statements after Discussion 7.5, Andrea showed interest in learning what statement would be obtained when negating a "some-statement". In this process she concluded that when negating a "some-statement", an "all-statement" will always result. For Andrea, this excluded the possibility of obtaining "no-statements"[204].

She then showed interest in finding out what needed to be negated to obtain a "no-statement". I drew the teachers' attention to the equivalence between "no-" and "all-" statements[205]. However, Andrea did not see the relevance of this link to determine what needed to be negated to obtain a "no-statement".

> Andrea: *Alright, but my question was, what should I negate in order to get "none"? I mean, when I negate "all", I get "some"; when I negate "some", I get "all"; but what should I negate in order to get "none"?*

Observe that Andrea had summarized the negations of "all-" and "some-" statements as:

- The negation of "all…" is "some…"
- The negation of "some…" is "all…"

She was concerned because she could not identify what needed to be negated to obtain a "none-statement".

Andrea's rejection of "none…" as the negation of "some…"

I drew the teachers' attention to the equivalent "all-statement" again and the universality of "no-statements". In particular, the teachers were asked whether the "no-statement" St108 was universal or existential.

> St108: *No odd number is an Innova number*

Andrea quickly answered that it was universal. Nevertheless, Andrea's impatience to know what statement should be negated in order to obtain "none…" was evident at that point.

[204] For details on this discussion, see Section II.3.1 above.
[205] For details on this issue, see Section III.1.1 above.

Andrea*: Sure, but what I am asking is, when I say, the negation of "some" is going to be "all", and what should I negate, what statement should I negate in order to get "none"? or there is not any?*

I: *Well, some…*

Andrea: *Some? (she looks surprised), no.*

(laughter)

Andrea was not satisfied with my answer (that a "no-statement" could be obtained when negating a "some-statement"). She rejected it (*"Some? [she looks surprised], no"*).

A plausible explanation for this phenomenon is that, as I explained before, Andrea was not willing to make a strong connection between "no-" and "all-" statements. Instead, she constantly relied on her initial interpretation of "no-statements".

The negation of "no-statements": A conflict

At this point of the intervention Andrea had acknowledged that St108 is equivalent to St109.

St108: *No odd number is an Innova number*

St109: *All odd numbers are not Innova numbers*

As a way to push Andrea to rely on the equivalent "all-statement" to negate a "no-statement" I asked her to negate St108, since we had not discussed yet the negation of "no-statements". My expectation was that once the teachers had negated the "all-statement" St109 and obtained a "some-statement", they could return to the "no-statement" St108 by negating the "some-statement", through double negation. This process revealed a new conflict.

Andrea provided statement St113 as the negation of St108 (she used her DSS-approach, see Section I.4).

St113: *Some odd numbers are not Innova numbers.*

St113 is not the negation of St108. Andrea avoided negating the equivalent "all-statement" and instead she negated St108 directly; otherwise, she would have obtained a different "some-statement". I then asked her to negate the equivalent "all-statement" St109. She provided St114 as the negation for St109, as expected,

St114: *Some odd numbers are Innova numbers.*

My expectation was that Andrea would become aware that the negation for St108 that she provided (St113) was not correct as the negation for St109 (St114) was different. Instead, Andrea claimed that *negating St108 was different from negating St109*.

Andrea's DSS-approach to negate "no-statements"

Andrea's negation of St108 was the first time she negated a "no-statement". Here I expand on the approach she used to do that.

Andrea formed her DSS-approach to cope with negations of universal and existential statements. The initials DSS stand for *Distribute, Separate and Substitute*, respectively, and they suggest the sequence of steps Andrea followed to negate statements[206].

As I explained above, Andrea's DSS-approach is based on the rules to negate universal and existential statements that I introduced during the intervention. Her DSS-approach was supported by Andrea's access to semantic substitutions that she used in the "substitute" step of the approach. For example, when she used her DSS-approach to negate "all-statements" of the form *"All X are Y"* she distributed the negator, separated the parts and substituted "not all" with "some". This means that as long as Andrea had a semantic substitution for the negation of the involved quantifier, she could use her DSS-approach to negate the statement.

When negating an affirmative "no-statement", Andrea assumed that the negation of "no/none" was "some" and used it for her substitution in her DSS-approach. In Figure 55 I show her DSS-approach applied to *"No X is Y"*, where I used the sign ~ to represent negation.

$$\begin{aligned} \sim(\textit{No } X \textit{ is } Y) &= \sim(\textit{No } X \textit{ is } Y) = \sim\!\textit{No } X \text{ is } \sim Y \\ &= [\sim\!\textit{No } X] \text{ is } [\sim Y] \\ &= \text{Some } X \text{ are not } Y \end{aligned}$$

Figure 55. Andrea's DSS-approach to negate "No X is Y"

Once she distributed the negator, she separated the statement to "prepare" it for the semantic substitution. The reason why Andrea negated "no X" as "some X" is not clear. It is possible that Andrea used her initial assumption that the negation of "no…" was "some…". This might be supported by the way Andrea disproved affirmative "no-statements", for which she assumed that *"No X is Y"* was false as she could find at least one *X* that was *Y*[207].

The conflict: The negation of an affirmative "no-statement"

Andrea's conflict emerged as she realized that her DSS-approach led to a negation of St108 that was different from what was expected.

> Andrea: *Alright, what if I negate the "no-statement", I mean, alright, I forget about this equivalent ["all-statement" St109] and I just want to negate that statement [St108]. I mean, I do not change it into its equivalent. I want to negate it [St108] directly… can I do that? ... How would that be?*
>
> I: *Look, if I have a statement of the form "No S is P" and you ask me for its negation, I would say its negation is "Some S are P" ...*
>
> Andrea: *Wouldn't it be "are NOT P"? ... so over there ["Some S are P"], the negation does not affect the second-, the consequent?*

---

[206] For details on Andrea's use of her DSS-approach for "all-", "some-" and "there is-" statements, see Section I.4 and Section II.3, respectively, above.

[207] For details on Andrea's disproving of affirmative "no-statements", see Section III.1.2 above.

The conflict consisted in the mismatch between Andrea's DSS-approach negation of *"No S is P"* and what was obtained when negating the equivalent "all-statement": *"Some S are not P"* and *"Some S are P"*, respectively. The incompatibility is evident in the different consequents of the two "some-statements". Andrea questioned that the consequent did not change in the expected negation (*"so over there, the negation does not affect the second-, the consequent?"*), which according to her DSS-approach should have changed to "not P" (*"Wouldn't it be "are NOT P"?"*).

Andrea's desire to avoid using the equivalent "all-statement" for St108 (*"I forget about this equivalent… I do not change it into its equivalent. I want to negate it [St108] directly"*) suggested again that Andrea did not get a real sense of the logical interpretation of "no-statements".

The conflict was finally "resolved" with Andrea's passive acceptance of the negation I suggested through the negation of the equivalent "all-statement". Andrea's negation of negative "no-statements" was not different from her negation of affirmative "no-statements".

In general, Andrea did not understand negation throughout the intervention. In my view, it is clear that Andrea did not establish an explicit relationship between the negation of a statement and its falsity, neither did she resort to any implicit linkage when negating statements. This is particularly true for the case of affirmative "no-statements", their negation and disproving. Andrea showed that she was aware of when an affirmative "no-statement" was false and why; however, when she was asked to find the negation of a "no-statement", she resorted to her DSS-approach, that resulted in an incorrect negation.

Summary of Section III.1

Andrea began the intervention with some initial assumptions related to the interpretation, negation and disproving of "no-statements". For Andrea, "no-statements" of the form "No X …" (both affirmative and negative) referred to the elements outside X. Her initial interpretation of "no-statements" was non-mathematical. Instead, it is closely related to the meaning of the quantifier "no/none" in everyday language in Spanish. Accordingly, Andrea assumed that the representation of "no-statements" (both affirmative and negative) was as two disjoint sets. Her disproving of affirmative "no-statements" was attuned with the mathematical perspective; she assumed that *"No X is Y"* was false if there was an X that was Y.

Because of her initial interpretation of "no-statements", Andrea did not link "no-" and "all-" statements, which made it difficult for her to understand the mathematical interpretation and negation of "no-statements". Notably, Andrea's DSS-approach to negate universal and existential statements added to the challenge as it led Andrea to experience a conflict when negating an affirmative "no-statement". Her DSS-approach involved a semantic substitution of the negation of "no" with "some". Andrea obtained a negative "some-statement" when she used her DSS-approach to negate an affirmative "no-statement", while she obtained an affirmative "some-statement" when she relied on the equivalent "all-statement".

Andrea's intervention-based equivalence for *"No X is Y"* (*"No X is not Y"*) as *"All X are not Y"* (*"All X are Y"*) and her negation of *"No X is Y"* as *"Some X are Y"* seems to have been of non-mathematical nature, but authoritarian, as she did not further challenge my inputs. Instead, she eventually accepted them. None of Andrea's assumptions about "no-statements" exhibited a real mathematical understanding of "no-statements".

Figure 56 shows the development of Andrea's assumptions related to "no-statements" and the elements or aspects of the intervention that might have supported or triggered changes[208].

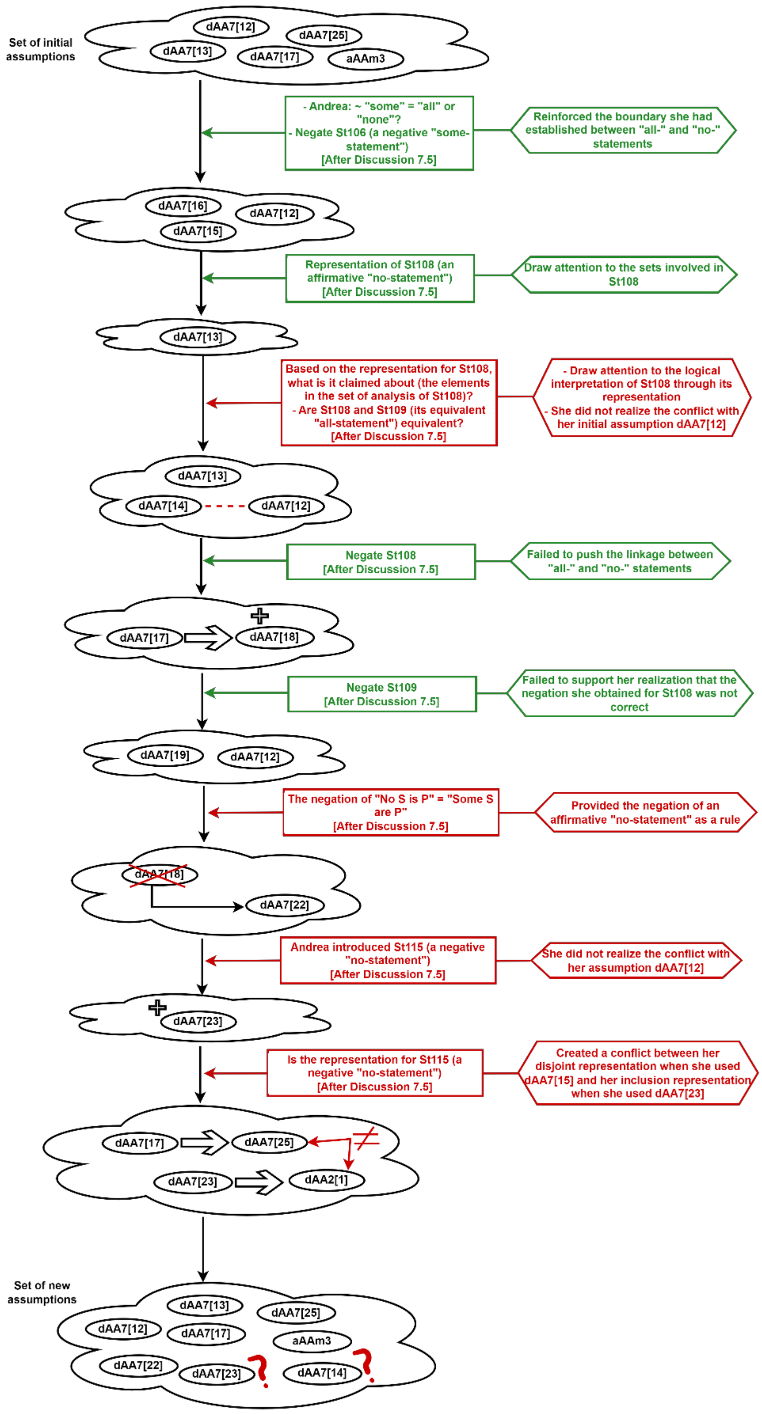


Figure 56. The development of Andrea's assumptions about "no-statements" (dAA7[12]: "No X is Y" and "No X is not Y" refers to everything that is not X; dAA7[13] : "No X is Y" and is represented as two disjoint sets X and Y; dAA7[17]: The negation of "no…" is "some…"; dAA7[25] : "No X is not Y" is represented as two disjoint sets X and Y; aAAm3: Given two disjoint sets X and Y, Y represents the set of the elements that are "not X"; dAA7[15]: The negation of a "some-statement" is an "all-statement", and never a "no-statement"; dAA7[16]: The negation of "Some X are Y" ("Some X are not Y") is "All X are not Y" ("All X are Y"); dAA7[14]: "No X is Y" is equivalent to "All X are not Y"; dAA7[18]: The negation of "No X is Y" is "Some X are not Y" (her DSS-approach); dAA7[19]: The negation of "All X are not Y" is "Some X are Y"; dAA7[22]: The negation of "No X is Y" is "Some X are Y"; dAA7[23]: "No X is not Y" is equivalent to "All X are Y"; dAA2[1]: "All X are Y" can be represented as X included in Y and Y − X may be empty)

[208] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

2. Universal Negative Statements with a different form from "no-statements"

In order to guide the teachers towards a mathematically-consistent interpretation of "no-statements", I addressed their equivalent "all-statements". As I explained before, Andrea did not associate "no-statements" with their equivalent "all-statements". That is the reason why I report here on universal negative statements (UNSs) that have a different form from "no-statements", for example negative "all-statements" (i.e., statements of the form *"All X are not Y"*). I explore here Andrea's initial interpretation of negative "all-statements", which might help understand her lack of association of "no-" and "all-" statements.

Two different forms of universal negative statements, other than "no-statements", were discussed before and during the intervention for teachers: (1) *"All X are not Y"* (negative "all-statements") and (2) *"For every x in N, P(x) is not Q"* (negative "for-every-statements"). In total we discussed four of these two types of UNSs, as listed in Table 30. While the negative "all-statements" St2 and St121 were provided to the teachers, St104 was introduced by Andrea and the negative "for-every-statement" St75 was formulated by the teachers as a conjecture.

Table 30. Universal Negative Statements analyzed before and during the intervention, their form and when they arose.

Code	Universal Negative Statement (UNS)	UNS form	When
St2	*All «Vallejo» numbers are not even numbers*	*(1) All X are not Y*	Before the Intervention
St75	*For every natural number* n *(different from zero), when substituted in the expression* $1 + 1141n^2$ *it is not a perfect square*	*(2) For every x in N, P(x) is not Q*	During the intervention (after D3.2)
St104	*All natural numbers divisible by 2 are not divisible by 3*	*(1) All X are not Y*	During the intervention (after D7.5)
St121	*All divisions of natural numbers do not have remainder zero*	*(1) All X are not Y*	During the intervention (D9)

I divide this section into two parts. The first part concerns negative "for-every-statements" (Section 2.1) and the second part is about negative "all-statements" (Section 2.2). I begin with negative "for-every-statements" given that there is only one aspect discussed about them during the intervention: their disproving. On the other hand, there are three aspects that were discussed about negative "all-statements": their interpretation (Section 2.2.1), their disproving (Section 2.2.2), and their negation (Section 2.2.3).

2.1. Negative "for-every-statements"

Andrea's disproving of a negative "for-every-statement"

Statement St75 (*For every natural number* n *(different from zero), the expression* $1 + 1141n^2$ *does not produce a perfect square*) was formulated by the teachers as a conjecture based on a pattern they identified after Discussion 3.2. Andrea tried to find a counterexample for it in an attempt to disprove it[209]. At that point, Andrea had become

[209] For details on Andrea's approach to solve the task and her "looking for a counterexample" approach, see Section I.2.2.2 above.

aware of the characteristics of counterexamples to affirmative "all-statements", which were tied to her mathematical interpretation of the statements. Andrea's attempt to disprove St75 included her search for a natural number for which $1 + 1141n^2$ was a perfect square number, that is, a counterexample to St75 (see Figure 57 for an abstract version of Andrea's disproving).

"For every x in N, P(x) is not Q" is false if there is an *x* in *N* such that *P(x)* is *Q*

Figure 57. Andrea's assumption dAA3(2) about disproving negative "for-every-statements".

Statement St75 was the only negative "for-every-statement" the teachers discussed during the intervention. It is likely that Andrea transferred her current understanding of disproving UASs to disproving St75. It is possible that the form of the statement supported this process. From Andrea's identification of suitable characteristics for a counterexample to St75, it is clear that her interpretation of St75 was mathematically aligned. This certainly played an important role when searching for a counterexample example with specific characteristics.

2.2. Negative "all-statements"

There are three ways in which Andrea's assumptions about negative "all-statements" were revealed: the way she interpreted them (Section 2.2.1), the way she disproved them (Section 2.2.2), and the way she negated them (Section 2.2.3).

2.2.1. Andrea's initial interpretation of negative "all-statements"

From a mathematical perspective, both universal negative statements *"All X are not Y"* and *"For every x in N, P(x) is not Q"*, despite their different forms, refer to all the elements in the set of analysis and have a negative consequent.

Unlike her interpretation of negative "for-every-statements", Andrea's initial interpretation of negative "all-statements" is non-mathematical. Her initial interpretation emerged before the intervention. Andrea assumed that:

> bAA8: *"All X are not Y"* is the same as *"Not all X are Y"*.

Andrea's assumption bAA8 arose during the First Exploratory Interview. One of the tasks asked the teachers to choose from a list of twelve statements those that conveyed the same as St2, *"All «Vallejo» numbers are even numbers"*[210]. Andrea showed hesitation about option "e" (*"The even numbers are not «Vallejo» numbers"*). She explained her rationale as follows:

> Andrea: *This [statement "e"] is OK, partially… **But all even numbers are not Vallejo [numbers]**… Of course, because here it only says that all Vallejo [numbers] are even [numbers], **but not all even numbers are Vallejo [numbers]**, … I understand that Vallejo [numbers] can be a group [she probably means a "set"] of numbers, but all this group of numbers are even. That is, it is a group of all even numbers, the Vallejo numbers is a subgroup [she probably means a "subset"]. Then… **the even numbers are not Vallejo numbers**, because when speaking about the even numbers I could refer to another [number] that is outside this subgroup.*

[210] Other aspects of Andrea's solution were discussed for example in Section I.1.1.1 above.

> *Therefore, in that case it does hold that-, the even numbers, those outside, are not Vallejo numbers.*

Based on these lines, I infer that Andrea assumed that the statements *"All even numbers are not Vallejo numbers"* and *"Not all even numbers are Vallejo numbers"* stated the same (Andrea's initial assumption bAA8). First, Andrea matched *"All even numbers are not Vallejo numbers"* with *"The even numbers are not «Vallejo» numbers"* (option "e"). Second, Andrea claimed that *"But all even numbers are not Vallejo [numbers]"* as she could infer this from St2. She then referred once more to St2 and what appears to be again a conclusion she made based on it (*"Of course, because here it only says that all Vallejo [numbers] are even [numbers], but not all even numbers are Vallejo [numbers]"*). It is as if she had arrived in both cases at the same conclusion but expressing it in different words. This is, I presume that Andrea treated both statements – *"All even numbers are not Vallejo numbers"* and *"Not all even numbers are Vallejo numbers"* – as interchangeable. Andrea's use of her assumption that a statement of the form *"The X are Y"* was equivalent to *"All X are Y"* was observed also later, during Discussion 1.5[211].

A possible explanation for Andrea's initial assumption bAA8

Andrea's initial assumption bAA8 could be explained in terms of the way simpler cases of negations are understood in ordinary language. For instance, in common language the sentences *"Alice is <u>not</u> responsible"* and *"<u>It is not true that</u> Alice is responsible"* make the same claim about Alice and whether she is responsible. In this example both negations, the *"internal"* (in the first statement) and the *"external"* (in the second statement) negations, do not change the interpretation of the sentence because of the position where the negator is placed (either it is in front of the statement, or it is after the verb[212]). Nonetheless, when this "flexibility" is used in a mathematical context, unfortunate conclusions can be drawn. For example, it might lead to claiming that *"All X are not Y"* and *"Not all X are Y"* are the same, which is clearly inconsistent with a mathematical perspective.

2.2.2. Andrea's disproving of negative "all-statements"

During the intervention Andrea attempted to prove a negative "all-statement" that was actually false. In her attempt, she revealed her lack of understanding of what was involved in disproving negative "all-statements".

Discussion 7.5 included a discussion about St103, its truth value and justification for it.

> St103: *Not all natural numbers divisible by 2 are divisible by 3.*

To analyze St103, Andrea had determined its equivalent statement *"Some natural numbers divisible by 2 are not divisible by 3"*. She proved St103 by proving this equivalent "some-statement". Andrea provided the number 8 to justify that St103 was true. A discussion developed about the negative "all-statement" St104 that Andrea proposed.

> St104: *All natural numbers divisible by 2 are not divisible by 3*

211 During Discussion 1.5 Andrea explained that the statement *"The people present in this classroom are minors"* referred to all of the people present in the classroom, not some or none, but all of them (see Section III.2.2.1 above).

212 In Spanish the negator comes before the verb.

Episode 18 includes the discussion we had about the truth value of St104 and its respective justification.

Episode 18

Turn	Who	What
1	Andrea	*That [St104] is true.*
2	Gessenia	*10*
3	I	*So, true or false?*
4	Both	*True.*
5	I	*Are you sure?*
6	Andrea	*Yes.*
7	I	*Then, if you say that this is true, what should we verify? What should we guarantee? What should we justify? What should my justification be like?*
8	Andrea	*An example.*
9	I	*An example?*
10	Gessenia	*Yes, one example is sufficient.*
11	I	*Which one is my set of analysis?*
12	Gessenia	*My antecedent?*
13	I	*Yes.*
14	Gessenia	*All natural numbers divisible by 2.*
15	I	*What should I verify for each of them? I say, all natural numbers divisible by 2. Check, check, check, … (I use gestures) … What should I verify for those?*
16	Gessenia	*That at least one is not divisible by 3.*
17	I	*Hmm, nope. If you say all of them, and you say that this is true, I should verify this for EACH OF THEM. I pick one even number, and it shouldn't be divisible by 3. Then another, and another. The verification is then for all the numbers divisible by 2. If you say that this is true, it would have to-, according to what you say, it [that it is not divisible by 3] must hold for every number divisible by 2.*
18	Andrea	*But it could not be either an even number.*
19	I	*Why not?*
20	Andrea	*Because if I say that it is even and I pick, hmm, number 6.*
21	I	*What about number 6?*
22	Andrea	*It is divisible by 3.*
23	I	*And?*
24	Gessenia	*10*
25	I	*Is that why this could not be false? Or, rather, that is why this is false?*
26	Andrea	*Hmm.*
27	I	*Let's see, 10? What about number 10?*
28	Gessenia	*It says, all natural numbers divisible by 2. One of them, I pick number 10.*
29	I	*Ok, I pick number 10, what about number 10?*
30	Gessenia	*It is not divisible by 3.*
31	I	*Quite the opposite, that [10] satisfies, it supports my statement, doesn't it? I mean, this is rather suggesting, this is supporting that it [the statement] may be true. But, does it suffice? I must keep analyzing and try to find a case that breaks the relation [in the statement]. Is there any [example] that breaks that relation?*
32	Gessenia	*12*
33	I	*What about 12?*
34	Gessenia	*12 is divisible by both numbers.*
35	I	*Then, what about that?*
36	Gessenia	*The statement-, is true. No, wait, it is false.*
37	I	*Is it false or true?*

38	Gessenia	*It is false because not all…*
39	I	*Indeed, it is false.*
40	Andrea	*Ohhhh!*

At first, Andrea seemed confident that the negative "all-statement" St104 *"All natural numbers divisible by 2 are not divisible by 3"* was true (turns 1–6). Furthermore, Andrea claimed that an example (for what she meant a *confirming* example) would prove that St104 was true (turns 7 and 8). Observe that a confirming example for St104 is an example of a number divisible by 2 that is not divisible by 3, which is the same type of evidence she used to prove St103 (*Not all natural numbers divisible by 2 are divisible by 3*). It is possible to argue that Andrea used an *empirical justification scheme* (Harel & Sowder, 1998) to conclude that St104 was true; however, at this point of the intervention Andrea had already modified her justification scheme[213]. Instead, I presume that Andrea's reasoning relied on her initial assumption bAA8 that *"All X are not Y"* is the same as *"Not all X are Y"*. According to it, because St104 (*"All natural numbers divisible by 2 are not divisible by 3"*) was the same as St103 (*"Not all natural numbers divisible by 2 are divisible by 3"*), then St104 was true because St103 was already proved to be true. Recall that Andrea proved St103 by proving its equivalent "some-statement": *"Some natural numbers divisible by 2 are not divisible by 3"*, for which she was aware that a confirming example was sufficient evidence. This I call Andrea's assumption dAA7[11].

dAA7[11]: *"All X are not Y"* is true because there is an *X* that is not *Y*

Andrea's assumption dAA7[11] was implicit and based on three of her current assumptions: (1) *"All X are not Y"* is the same as *"Not all X are Y"* (Assumption bAA8), (2) the negation of *"All X are Y"* is *"Some X are not Y"* (Assumption dAA7[10]), and (3) one confirming example is sufficient to prove a "some-statement" (Assumption dAA7[4]). Because it was based on one of her initial assumptions (bAA8) and two of her intervention-based assumptions (dAA7[10] and dAA7[4]), I consider that her assumption dAA7[11] stemmed from the intervention.

A shift from the teachers' initial approach to prove St104 took place as I drew their attention to the logical interpretation of the statement and, moreover, the set of analysis in St104. I asked the teachers what each of the elements therein needed to be in order to prove that the statement was true (turns 11, 15 and 17). Gessenia properly identified that the set of analysis of St104 was the set of numbers divisible by 2, though she still assumed that a confirming example sufficed to prove St104 (turn 16). While Gessenia was still focused on an example of a number divisible by 2 and not divisible by 3 (the number 10), Andrea seemed conflicted by the number 6. This example puzzled her as she noticed that it was indeed an element in the set of analysis, and that it satisfied the second condition (6 is divisible by 3) (turns 18–26). The existence of a number with such characteristics did not seem to have been easily recognized by Andrea as enough evidence to disprove the statement[214]. Nevertheless, the rest of the discussion might have been what Andrea needed to fill in the gaps of her ongoing reasoning. Gessenia played a leading role and provided an example that supported St104 (the number 10). I highlighted the supporting nature of Gessenia's example and asked for the sufficiency of such examples to prove St104. Then I asked her to consider whether she could find a case that would "break" the relation in St104 (turn 31). Gessenia provided the number 12 and explicitly stated that her

[213] For details on Andrea's understanding of the status of confirming examples when proving USs, see Section I.3.1.1 above.

[214] At this point of the intervention, though, Andrea was aware of when a UAS was false and the characteristics of the counterexample to a UAS (see Section I.2.1.2 and Section I.2.2.2).

example satisfied both conditions (it is divisible by 2 and by 3) (turn 34), which meant that the statement was actually false (turn 36) as it did not hold for all cases. Andrea's surprise was evident (turn 40), presumably as she did not expect St104 to be false.

Andrea's new insight for the disproving of negative "all-statements" was explicitly revealed later in the intervention, where her focus was on the mathematical interpretation of the statement. During Discussion 9 Andrea used her emerging understanding to disprove St121 (*"All divisions of natural numbers do not have remainder zero"*): her new assumption dAA9[1],

> dAA9[1]: *"All X are not Y"* is false because there is an example of *X* that is *Y* (a counterexample)

Andrea provided a suitable counterexample to St121, which showed that she had already changed her initial assumption bAA8.

The role of the form of the statements

Even though the four statements in Table 30 are all universal negative statements, their form is not strictly the same. This difference in form had an effect in Andrea's initial interpretation of the statements.

In this section I showed that as a result of combining Andrea's initial assumption bAA8 (*"All X are not Y"* ≡ *"Not all X are Y"*) with her assumption that *"Not all X are Y"* ≡ *"Some X are not Y"*[215], Andrea seems to have initially interpreted the UNS *"All X are not Y"* as an existential statement. Evidence of Andrea's assumption was exhibited through her evaluation of a (false) UNS, which she assumed to be true because there was a confirming example for it.

In contrast, Andrea's evaluation of a (false) negative "for-every-statement" showed that Andrea was clearer about its structure and therefore her interpretation of it. She seemed confident of what to aim at when (attempting to) disproving the statement. This involved that Andrea correctly identified the characteristics that possible counterexamples should satisfy to disprove the negative "for-every-statement".

2.2.3. Andrea's negation of negative "all-statements"

Here I show that Andrea's negation of negative "all-statements" involved her use of her *Distribute, Separate and Substitute* (DSS-) approach to negate statements.

Andrea's DSS-approach to negate negative "all-statements"

Andrea exhibited the use of her DSS-approach[216] to negate negative "all-statements" after Discussion 7.5. Without hesitation Andrea concluded that the negation of St121:

> St121: *All divisions of natural numbers do not have remainder zero*,

was the statement *"Some divisions of natural numbers have remainder zero"*.

[215] See Andrea's negation of UAS in Section I.4 above.

[216] For details on Andrea's DSS-approach, see e.g., Section I.4 above.

Andrea's DSS-approach to negate negative "all-statements" is shown in abstract terms in Figure 58.

~ (*All X are not Y*) = ~ (*All X*) *are* ~ *(not Y)* = *[~ (All X)] are [~ (not Y)]* = *Some X are Y*

Figure 58. Andrea's DSS-approach to negate statements of the form "All X are not Y"

Her DSS-approach consists of distributing the negation, separating the statement to prepare it for substitution, substituting the negation of "all X" ("not all X") with "some X" and switching "are not Y" to "are Y".

Andrea's use of the DSS-approach for the case of negative "all-statements" was an extension of the DSS-approach she used to negate affirmative "all-statements".

Summary of Section III.2

Among the universal negative statements with a form different from "no-statements", the teachers and I discussed one negative "for-every-statement" (*"For every x in N, P(x) is not Q"*) and three negative "all-statements" (*"All X are not Y"*). Andrea's initial interpretation of those two types of statements differed. Unlike her initial interpretation of negative "for-every-statements", her initial interpretation of negative "all-statements" was not mathematical. Andrea's disproving of the negative "for-every-statement" during the intervention gave access to her interpretation of that type of statement. This was shown by her correct identification of the characteristics of counterexamples to the statement. In contrast, her initial interpretation of negative "all-statements" was observed before the intervention. Andrea equated the statements *"All X are not Y"* and *"Not all X are Y"*, which prevented her from seeing that a negative "all-statement" was universal. Her assumption that *"Not all X are Y"* was the same as *"Some X are not Y"* led her to assume that a negative "all-statement" was existential. This is the main reason why Andrea tried to prove a negative "all-statement" that was actually false. Given that she was aware that there was a confirming example for the statement *"Some X are not Y"* (equivalent to *"Not all X are Y"*), she assumed that *"All X are not Y"* was true. Once her attention was drawn to the mathematical interpretation of the statement, Andrea understood that the negative "all-statements" were universal and that to disprove them it is necessary to find a counterexample.

Andrea used her DSS-approach, which she developed during the intervention, to negate negative "all-statements". Even though this approach led her to obtain mathematically-consistent negations of negative "all-statements", it was a shortcut that did not involve mathematical understanding. I do not have evidence of how she would have negated negative "for-every-statements".

Figure 59 shows the development of Andrea's assumptions related to universal negative statements with a different form from "no-statements" and the elements or aspects of the intervention that might have supported or triggered changes[217].

[217] See Figure 14 (above) for an explanation of the shapes and colors used in this diagram.

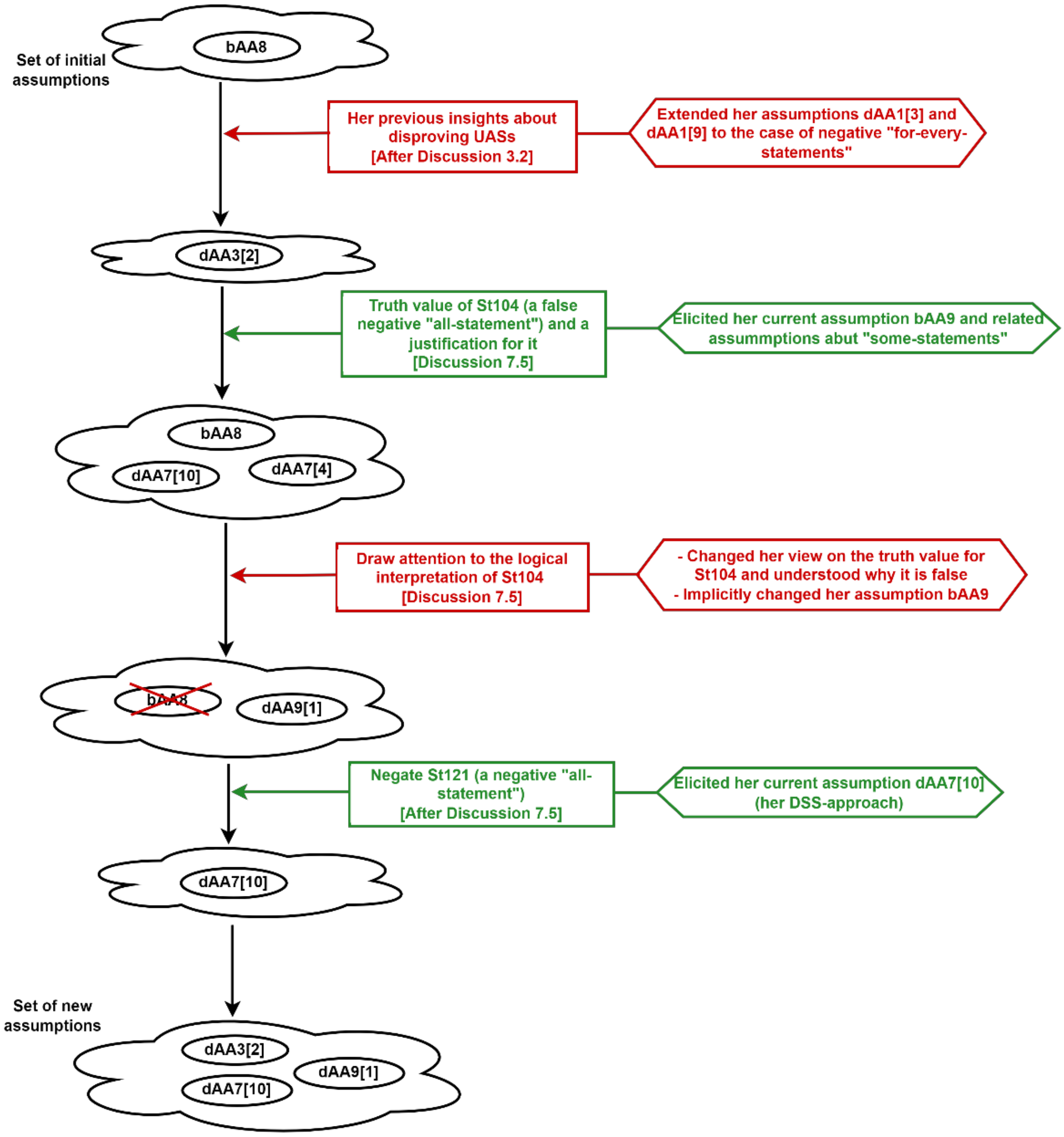


Figure 59. The development of Andrea's assumptions related to universal negative statements with a different form from "no-statements" (bAA8: "All X are not Y" is the same as "Not all X are Y"; dAA3[2]: "For every x in N, P(x) is not Q" is false because there is an x in N such that P(x) is Q; dAA7[10]: The negation of "All X are Y" is "Some X are not Y" ("Some X are Y"); dAA7[4]: One confirming example is sufficient to prove a "some-statement"; dAA9[1]: "All X are not Y" is false because there is an X that is Y)

Chapter 6: Conclusions

In this chapter I answer three of my four research questions:

- RQ1: How do in-service primary school teachers' assumptions related to dis/proving change while engaged in an intervention focused on Proof-Based Teaching (PfBT) and understanding the nature of proving?
- RQ1a: What features of the intervention led to the observed changes?
- RQ2: How are the in-service primary school teachers' assumptions that changed visible during their teaching in schools?

I divide this chapter into three main sections, according to each of three of my research questions. Section I has a focus on the conclusions related to the assumptions the teachers used and whether they changed or not during or after the intervention (RQ1). Section II draws attention to the conclusions related to the features of the 2018-intervention that might explain why changes were fostered or hindered (RQ1a). Section III has a focus on the conclusions about the ways the teachers' assumptions that changed were visible during their teaching (RQ2). RQ3 is the topic of Chapter 7.

I. Conclusions about the proof-related assumptions the teachers used

During the 2018-intervention the teachers exhibited, developed, changed and adapted many proof-related assumptions. Some of them were mathematical and some were non-mathematical. A *mathematical assumption* is an assumption that is mathematically aligned as it is formulated. An example of a mathematical assumption is: the negation of *"All X are Y"* is *"Some X are not Y"*.

Here I include conclusions concerning the teachers' assumptions, in six sections: unusual assumptions (Section 1), assumptions that reveal connections between statements (Section 2), assumptions that reveal that equivalent statements are treated differently (Section 3), mathematical assumptions based on non-mathematical reasons (Section 4), assumptions that involve the use of mathematical terms with non-mathematical meanings (Section 5) and assumptions that are a result of overgeneralizations (Section 6).

1. Unusual assumptions

Most of the non-mathematical assumptions the teachers used have been documented in the literature and seem to be fairly common among teachers as well (see Chapter 2, Section I). For example, the assumption that the converse of a true UAS is true[218] (e.g., Epp, 1999, 2003; Hoyles & Küchemann, 2002; Wason, 1968)[219] and the assumption that the statements *"Not all X are Y"* and *"No X is Y"* are equivalent[220] (e.g., Pasztor & Alacaci, 2005; Epp, 1999). However, during the intervention I also observed some unusual initial assumptions Andrea began the intervention with. For example, she assumed that:

- The converse of a true universal statement is false[221] (bAA2)

[218] Gessenia and Lizbeth used this (bAG1) assumption (for details, see Chapter 5, Section I.1.1.2).
[219] See Chapter 2, Section I.1, for a review on the "converse error", possible explanations why this assumption might emerge and hints for possible solutions based on previous research.
[220] None of the three teachers who participated in the 2018-intervention expressly used this assumption.
[221] For details, see Chapter 5, Section I.1.1.1.

- *"Not all X..."* is different from *"No X..."*[222] (dAA7[2])

While Andrea's unusual assumption bAA2 was non-mathematical, her assumption dAA7[2] was mathematical. During the intervention, her assumption bAA2 changed into the following mathematical assumption:

- If a UAS is (assumed to be) true, its converse *may* be false[223] (aAAm2)

The change of her initial assumption bAA2 was revealed through her use of more precise terms. Instead of claiming that the converse *is* false, Andrea was more cautious and claimed that it *may* be false.

The two assumptions bAA2 and dAA7[2] do not seem to have been reported in the literature as used by teachers when engaged in proof-related discussions. In fact, Hoyles and Küchemann (2002) found out that high-attaining high school students assumed that a conditional statement and its converse have different truth values. Nevertheless, no study that came to my attention have reported that in- or pre-service teachers share a similar assumption. It is important to become aware of the teachers' unusual proof-related assumptions to be prepared to tackle them when needed.

2. Assumptions that reveal connections between statements

During the intervention Andrea used assumptions that revealed connections she had made between statements. For example, Andrea related "all-statements" and "some-statements". This was revealed with her use of the following assumptions:

- *"Not all X are Y"* is the same as *"Some X are Y"*[224] (dAA7[1])
- If *"All X are Y"* is true, then *"Some X are Y"* is false[225] (dAA1[1])
- If *"Some X are Y"* is true, then *"Some X are not Y"* is true[226] (dAA7[3])
- "Some" means "from all, one group, but not all"[227] (dAA1[8])

Andrea assumed that the simple implicit negation of an affirmative "all-statement" was an affirmative "some-statement" (assumption dAA7[1]). She relied on her semantic substitution of "not all" with "some" to conclude dAA7[1]. For Andrea affirmative "all-statements" and "some-statements" could not be both true or false; if one of the statements was true, the other was false (assumption dAA1[1]). She assumed that both affirmative and negative "some-statements" were true if one of them was true (assumption dAA7[3]). Her three first initial assumptions (dAA7[1], dAA1[1] and dAA7[3]) were based on Andrea's initial non-mathematical interpretation of "some" (assumption dAA1[8]). Andrea had established a connection between her "all-" and "some-" statements because of her assumption dAA1[8], which showed to be fundamental to the other assumptions.

During the intervention Andrea accepted the initial meaning input for "some" as "at least one", which did not involve that she had abandoned her assumption dAA1[8]. She changed her fundamental assumption dAA7[3] first, which had an effect in all her related assumptions that changed into:

[222] For details, see Chapter 5, Section I.4.
[223] For details, see Chapter 5, Section I.1.1.1.
[224] For details, see Chapter 5, Section I.4.
[225] For details, see Chapter 5, Section II.2.
[226] For details, see Chapter 5, Section II.1.
[227] For details, see Chapter 5, Section II.1.

- The negation of *"All X are Y"* is *"Some X are not Y"*[228] (dAA7[10])
- If *"Some X are Y"* and *"All X are Y"* are both true, then *"Some X are not Y"* is false[229] (dAA7[8])
- If *"Some X are Y"* is true and *"All X are Y"* is false, then *"Some X are not Y"* is true (dAA7[9])
- "Some" means "at least one, and maybe all" (dAA7[7])

Andrea's new assumptions were mathematical; however, her assumption dAA7[10] was grounded on a non-mathematical form of reasoning. Below in Section I.4 I refer back to this point.

Another example is the connection Andrea made between "there-exist-statements" and "no-statements". This was observed with her use of the following assumption:

- *"There does not exist X that is Y"* is the same as *"No X is Y"*[230] (dAA9[4])

The connection Andrea made between "there-exist-" and "no-" statements was through negation. She assumed that the negation of a "there-exist-statement" is a "no-statement". To find an equivalent statement for the affirmative "there-exist-statement" she used her semantic substitution of "there does not exist" with "no" (her Separate and Substitute approach). Andrea's mathematical assumption dAA9[4] did not change during the intervention. A similar assumption was challenged during Discussion 9. Andrea found two different equivalent statements for the simple implicit negation of a negative "there-exist-statement", *"There does not exist X that is not Y"*.

- *"There does not exist X that is not Y"* is the same as *"No X is not Y"*[231] (dAA9[6])
- *"There does not exist X that is not Y"* is the same as *"No X is Y"*[232] (dAA9[8])

With her Separate and Substitute approach she obtained her assumption dAA9[6], while as Andrea distributed the negator, separated the negated parts and substituted "there does not exist" with "no", she obtained her assumption dAA9[8]. This I called her Distribute, Separate and Substitute (DSS-) approach[233] to negate existential statements. Her hesitation exhibited Andrea's lack of understanding of the negation of single-quantified (SQ-) statements.

This shows the importance of becoming aware of the way the teachers relate SQ-statements. As I have shown in my work, statements are not isolated entities, neither are the teachers' assumptions about them. Accessing a broad range of the teachers' assumptions can support identifying the connections the teachers make between statements and determine the root obstacles the teachers have when making mathematically aligned proof-related assumptions. Teachers might use these connections implicitly. Having explicit discussions about them facilitates the access to the assumptions that reveal those links.

[228] For details, see Chapter 5, Section I.4.
[229] For details, see Chapter 5, Section II.1.
[230] For details, see Chapter 5, Section II.3.2.
[231] For details, see Chapter 5, Section II.3.2.
[232] For details, see Chapter 5, Section II.3.2.
[233] For details, see Chapter 5, Section II.3.

3. Assumptions that reveal that equivalent statements are treated differently

Equivalent SQ-statements can have different forms of expression (e.g., see Chapter 3, Section II.2). For instance, existential statements can be presented as "some-statements", "there-is-statements" and "there-exist-statements". However, it is possible that a lack of awareness of this link is revealed through the assumptions that are used. For example, Andrea distinguished between equivalent "some-statements" and "there-is-statements", which was observed through her use of the following assumptions:

- If *"All X are Y"* is true, then *"Some X are Y"* is false[234] (dAA1[1])
- If *"All X are Y"* is true, then *"There is X that is Y"* is true[235] (bAA7)

While her assumption dAA1[1] is non-mathematical, her assumption bAA7 is. Andrea's assumptions revealed that she treated "some-statements" and "there-is-statements" differently; otherwise, she would have concluded the same truth value for both equivalent existential statements. This suggests that some teachers might not find obvious that "some-statements" and "there-is-statements" are two different in form but equivalent existential statements.

Andrea's assumption dAA1[1] changed as Andrea learned that a true "some-statement" might be universally true, which she initially rejected (see Section I.2).

Another example is Andrea's rejection of the equivalence between "no-statements" and negative "all-statements". She initially assumed that:

- *"No X is Y"* is different from *"All X are not Y"*[236] (dAA7[20])

Andrea avoided relying on equivalent negative "all-statements" when representing and negating "no-statements" (for details, see Chapter 5, Sections III.1.1 and III.1.3). During the intervention she accepted the equivalence between the statements *"No X is Y"* and *"All X are not Y"*; however, her agreement was not based on any form of understanding as she continued to use her assumption dAA7[20] afterwards. This shows that Andrea did not change her initial assumption dAA7[20]. The root problem with it was Andrea's initial interpretation of the quantifier "no/none", which I further discuss in Section I.4.

A third example for this section is the distinction Andrea made between negative "all-statements" and negative "for-every-statements" when disproving them. Andrea assumed that a statement of the form *"For every x in N, P(x) is not Q"* was false if there was an *x* in *N* such that *P(x)* was *Q*. She used this assumption in her attempt to find out the truth value of a (true) negative "for-every-statement" as she tried to find a counterexample for it. However, when discussing a (false) negative "all-statement", Andrea did not even consider the possibility for it to be false, which indicates that she did not make connections between negative "for-every-statements" and negative "all-statements" (for details, see Chapter 5, Sections III.2.1 and III.2.2).

Andrea's assumption for disproving negative "for-every-statements" is intervention-based as it was influenced by previous discussions on the characterization of counterexamples to universal affirmative statements. On the other hand, her assumption about the truth value of negative "all-statements" was first influenced by one of her non-mathematical initial assumptions. Andrea assumed that *"All X are not Y"* is the same as

[234] For details, see Chapter 5, Section II.2.
[235] For details, see Chapter 5, Section II.1.
[236] For details, see Chapter 5, Section III.1.1.

"Not all X are Y", which ultimately turned her attention to proving its equivalent "some-statement" *"Some X are not Y"*.

Learning about these assumptions shows the importance of discussing SQ-statements in different forms of expression as the teachers might not see that they are equivalent. These examples show that even though statements might be logically equivalent, they might not be perceived as such. Attention needs to be paid to the way statements are expressed as teachers might not make connections between statements with the same logical interpretation. Research that explores the way in-service teachers interpret equivalent SQ-statements before engaging them in proof-related discussions about them does not seem to have been developed yet.

4. Mathematical assumptions based on non-mathematical reasons

The fact that the teachers use mathematical assumptions do not necessarily mean that the rationale behind those assumptions is mathematical. For example, Andrea's initial interpretation of "no-statements" was non-mathematical. She assumed that:

- *"No X..."* refers to everything that is not *X*[237] (dAA7[12])

Nevertheless, she used assumptions about "no-statements" that were mathematically aligned. For example, she assumed that:

- *"No X is Y"* is represented as two disjoint sets *X* and *Y*[238] (dAA7[13])
- The negation of *"No X is Y"* is *"Some X are Y"*[239] (dAA7[22])

Andrea's assumptions about the representation of affirmative "no-statements" (assumption dAA7[13]) and the negation of affirmative "no-statements" (assumption dAA7[22]) were mathematical even though she reasoned from her non-mathematical interpretation of "no-statements" (assumption dAA7[12]). Her three initial assumptions did not change during the intervention. More importantly, her initial assumption dAA7[12] was resistant to change as I observed through her rejection of the equivalence with "all-statements" (assumption dAA7[20]) and her permanent reliance on dAA7[12] to reason about "no-statements".

A second example of this conclusion is Andrea's negation of "all-statements" after the intervention. Her intervention-based assumption that:

- The negation of *"All X are Y"* is *"Some X are not Y"*[240] (dAA7[10])

involved a non-mathematical procedure to obtain a mathematical negation of "all-statements". Andrea used her DSS-approach to negate universal statements: she distributed the negator, separated the negated parts and substituted "not all" with "some"[241]. With her non-mathematical approach Andrea managed to obtain mathematical negations of "all-statements". This was Andrea's last assumption related to the negation of "all-statements", as she began the intervention with her assumption dAA7[1] that the negation of *"All X are Y"* is *"Some X are Y"*.

[237] For details, see Chapter 5, Section III.1.1.
[238] For details, see Chapter 5, Section III.1.1.
[239] For details, see Chapter 5, Section III.1.3.
[240] For details, see Chapter 5, Section I.4.
[241] For details, see Chapter 5, Sections I.4.

A third example is Andrea's initial assumption about whether a UAS and its converse state the same thing. She assumed that:

- A UAS and its converse do not state the same thing[242] (bAA3)

Andrea's assumption bAA3 is mathematically aligned; however, it was based on her non-mathematical initial assumption that:

- If a UAS is (assumed to be) true, its converse is false (bAA2)

Her assumption bAA3 was grounded on the different truth values that Andrea assigned to a UAS and its converse.

Her assumption bAA2 was also based on one of her non-mathematical initial assumptions:

- *"All X are Y"* can be represented as set X included in set Y, with $Y - X \neq \emptyset$ (bAA1)

Andrea's initial assumption bAA3 did not change, but Andrea found a new argument for it that was grounded on concepts that were introduced during the intervention. Andrea explained that a UAS and its converse do not state the same thing because the "sets of analysis" of the statements were not the same. On the other hand, her initial assumptions bAA2 and bAA1 changed, respectively, to the assumptions:

- If a UAS is (assumed to be) true, its converse *may* be false[243]
- *"All X are Y"* can be represented as X included in Y and $Y - X$ may be empty[244]

Learning about these cases is important as it shows that teachers might reach mathematically accurate conclusions; however, they might be based on rules and/or interpretations of non-mathematical nature. Reporting results of research where the development of proof-related assumptions is the focus should involve investigating where those assumptions stem from. Exploring the teachers' reasons that support the proof-related assumptions they hold has received not enough attention in the existing literature. Not being aware of the origins of assumptions might create the false idea that logical principles are understood when they are not. For example, in many studies that have a focus on pre- and in-service teachers' justifications, high percentages of the participants produced valid proofs or evaluated arguments for universal or existential arithmetic statements with mathematically aligned criteria (e.g., Barkai et al.'s, 2002; Tabach et al., 2010a; Tabach et al. 2010b); however, it is unknown whether the assumptions the teachers used for the proofs they produced were mathematically grounded.

5. Assumptions that involve the use of mathematical terms with non-mathematical meanings

The teachers used mathematical terms with non-mathematical meanings to express their proof-related assumptions. For example, at the beginning of the 2018-intervention Lizbeth used the term "justify" with a non-mathematical meaning for it. She used this term and its non-mathematical meaning in her assumptions that:

[242] For details, see Chapter 5, Section I.1.1.1.
[243] For details, see Chapter 5, Section I.1.1.1.
[244] For details, see Chapter 5, Section I.1.1.1.

- Confirming examples can *justify* that a (true) UAS that involves infinite cases is true[245] (bAL1)

Much later during the intervention she began to shape her assumption bAL1 as new terms were introduced during the intervention and she gained insights about the status of confirming examples when proving USs. Lizbeth assumed that:

- Confirming examples *justify* that a false UAS that admits confirming examples with infinite cases involved is true, though they do not necessarily *guarantee* that the UAS is true (dAL1[3])

Her new assumption dAL1[3] shows that Lizbeth used the term "justify" to mean "confirm, verify, support", but not anymore as "sufficient to show that a US is true". The introduction of the term "guarantee" allowed her to make that distinction. She assumed that verifying some examples does not *guarantee* that a UAS is true. All examples need to be verified in order to do that (her assumption dAL3[1] in Chapter 5, Section I.3.1.2).

Assumption dAL1[3] changed as Lizbeth adapted her use of the term "justify" to mean what she meant by "guarantee". This was revealed through her new assumption:

- In order to *justify* that a statement is true, we should verify each case involved in the statement (dAL3[2])

A second example of this conclusion is Andrea's use of the terms "logically" and "deduce". Andrea used the expression *"mathematically not, but logically, yes"* to explain that from a mathematical perspective *"Some X are Y"* did not imply that *"Some X are not Y"*, but from a "logical" point of view, it did[246]. With the term "logical" Andrea meant what she would have expected from her everyday reasoning or common sense, in contrast to what was expected from a mathematical view. In the same context Andrea claimed: *"I can deduce one [statement] from the other one"*. Andrea's assertion suggested again her personal certainty about her initial assumption that *"Some X are Y"* implies that *"Some X are not Y"*. Andrea's use of the terms *"logical"* and *"deduce"* were not of mathematical nature.

A third example is Andrea's initial use of the quantifiers "some" and "no". She assumed that "some" means "from all, one group, but not all" (assumption dAA1[8]) and assumed, for example, that: *"Not all X are Y"* is the same as *"Some X are Y"* (assumption dAA7[1]); if *"All X are Y"* is true, then *"Some X are Y"* is false (assumption dAA1[1]); if *"Some X are Y"* is true, then *"Some X are not Y"* is true (assumption dAA7[3]). In the same vein, she assumed that *"no X..."* refers to everything that is not *X* (assumption dAA7[12]). Based on this meaning of "no" she assumed that *"No X is Y"* is represented as two disjoint sets *X* and *Y* (assumption dAA7[13]), and the negation of *"No X is Y"* is *"Some X are Y"* (assumption dAA7[22]).

In every conversation it is important to be aware of the meaning of the terms used so that understanding each other is achieved. Mathematical discussions are no exception. This conclusion is important as it supports other researchers' observations that the everyday context is in most cases the first source of reference that teachers and others use for the terms they use to express her mathematical reasoning (e.g., Halliday, 1978; Lee & Smith, 2009; Tall, 1977). In a recent study, Hempel & Buchbinder (2022) have found out that prospective secondary teachers use everyday language words (e.g., "opposite") when engaged in indirect proving. In general, the incompatibilities between the teachers' initial

[245] For details, see Chapter 5, Section I.3.1.2.
[246] For details, see Chapter 5, Section II.1.

assumptions and related mathematical assumptions can be accounted for the different interpretations of key terms. Although starting proof-related discussions from everyday statements with similar interpretations to mathematical ones and then moving to mathematical statements might be useful (e.g., Durand-Guerrier et al., 2012; Epp, 2003), it may also introduce potential obstacles as I have shown in my work. The teachers assumed that their initial everyday interpretation of universal and existential quantifiers could be applied to the mathematical context without any further discussion. That is, it might invite teachers to rely on their daily-life forms of reasoning and not to switch to the mathematical register. It is crucial to identify the meanings of the terms used and find ways to support the transit of meanings from their *daily-life register* to the *mathematics register* (Halliday, 1978).

6. Assumptions that are a result of overgeneralizations

The teachers tended to overgeneralize their observations and make them new assumptions during the intervention. For example, Gessenia tried to overgeneralize that the "set of analysis" could be always located to the front of a statement. She was influenced by the examples we analyzed, where the sets of analysis were the first sets in the statement (for details, see Chapter 5, Section I.1.1.2).

Gessenia also overgeneralized the form of expression in which justifications should be written. Based on the text-form of the justifications previously discussed during the intervention, Gessenia overgeneralized that all justifications should have a narrative format. Her assumption was revealed as she evaluated Lizbeth's justification for a false US, which included a numerical counterexample. Gessenia seemed conflicted and what she said revealed the assumption she had made: "*As we have always justified with text. But can a justification include an exercise?*" (for details, see Chapter 5, Section I.2.1.1).

Andrea also made some overgeneralizations. The teachers' recent insights about disproving USs had an impact on Andrea's assumptions about disproving ESs. She thought that "counterexamples" could also disprove "some-statements". Andrea assumed that:

- A "some-statement" is false if it has a counterexample[247] (dAA7[5a] and dAA7[5b])

Andrea acknowledged that she was unsure about her intervention-based assumptions dAA7[5a] and dAA7[5b] because we had not discussed false "some-statements" before. As Andrea's attention turned to the logical interpretation of "some-statements" and the evaluation of the possibility for them to be true, Andrea's assumption changed to:

- *"Some X are not Y"* is false if it is impossible to find an example of *X* that is not *Y*[248] (dAA7[6])
- *"Some X are Y"* is false if it is impossible to find an example of *X* that is *Y*[249](aAAt1)

The development of Andrea's assumptions about the status of confirming examples when proving USs also includes an overgeneralization Andrea made. It involved her transit from her assumption:

[247] For details, see Chapter 5, Section II.2.
[248] For details, see Chapter 5, Section II.2.
[249] For details, see Chapter 5, Section II.2.

- Confirming examples are sufficient to prove a UASs that involve infinite cases[250] (bAA5)

to her overgeneralization that:

- Confirming examples are insufficient to prove a statement (dAA1[10])

Andrea changed her assumption dAA1[10] through a series of several refinements:

- Confirming examples are insufficient to prove a universal statement[251] (dAA1[10]*)
- As long as the set of analysis is large, examples are not valid; if it is a small set, then they are[252] (dAA3[1])
- When the statement is universal and true with infinite cases involved, an example is not a valid justification, unless it is a generic example[253] (dAA9[9])

Generalizing is not a rare behavior in everyday life. We are usually generalizing things from our observations. Furthermore, this is similar to the way sciences different from mathematics build theories: they make observations, make a conjecture that is usually a generalization from previous observations, test the conjecture and if new data confirm the conjecture, then a conclusion or theory is drawn, which is normally submitted to revision if it is warranted with evidence. This is how scientific, non-mathematical, theories are "proved" or evidenced. However, mathematical theories or theorems involve proofs of different nature; as Krantz (2011)[254] points out, "in mathematics, once correct is always correct" (p. 224). There is nothing wrong with generalizing, but it becomes an issue when it goes too far towards overgeneralizing. The proofs of mathematics keep us from overgeneralizing.

All the teachers' overgeneralizations during the 2018-intervention were intervention-based assumptions and as such it is important to be aware that they might arise and need to be addressed. The fact that the teachers tend to overgeneralize regularities they observe is a sign that proof-related discussions should avoid suggesting patterns that invite to overgeneralizations. If overgeneralizations are made, discussions about them should not be avoided; otherwise, they can become obstacles for other discussions.

7. Assumptions that reveal the emergent development of further forms of reasoning

There were three interesting assumptions that Andrea and Lizbeth began to develop during the intervention. They emerged as a result of other well-settled assumptions and are as follows:

- If we do not find any counterexamples for the US conjecture, then it is a mathematical truth (dAL3[3])

[250] For details, see Chapter 5, Section I.3.1.1.
[251] For details, see Chapter 5, Section I.3.1.1.
[252] For details, see Chapter 5, Section I.3.1.1.
[253] For details, see Chapter 5, Section I.3.1.1.
[254] In fact, Krantz (2011) claims that the nature of proof should change and develop. For example, he discusses "computer proofs", "proofs by way of physical experiment" and "proofs by way of numerical calculation".

- *"Some X are not Y"* is false because it is impossible to find an example of *X* that is not *Y* and *"Some X are Y"* is false because it is impossible to find an example of *X* that is *Y* (dAA7[6] and aAAt1)
- A verbalized semi-general[255] counterexample is valid to disprove a US (aAAm1)

The consideration of *impossible* counterexamples[256] was particularly important for Lizbeth's assumption about proving UASs as eliminating the possibility of the existence of counterexamples. Her assumption dAL3[3] was a spontaneous result of our previous discussions about two main topics: (1) *disproving UAS* (the sufficient evidence to disprove a UAS and the description of counterexamples) and (2) *the status of confirming examples when proving UASs*. This case is important as Lizbeth brought together her assumptions related to disproving and proving universal statements. It shows the link between proving and disproving universal statements and the importance of understanding both processes conjointly. In this context, the main core of Lizbeth's assumption is the idea that for a US to be true, it is impossible that it is false. That is, it means that it is impossible to find counterexamples, which would need to be shown in order to guarantee the truth of the statement (for details, see Chapter 5, Section I.3.3).

Yopp (2017) has discussed the importance of a conceptualization of proof as eliminating counterexamples as a way to develop indirect reasoning. In his research, Yopp claims that "eliminating counterexamples can be viewed as a result (and the goal) of any proof of a generalization as well as a process for constructing a proof" (p. 155). Lizbeth's assumption dAL3[3] is similar to a view of eliminating counterexamples as a process for constructing a proof. She seems to understand the process of eliminating counterexamples as a guarantee to conclude that a US-conjecture is a mathematical truth, though it does not necessarily mean that Lizbeth actively used this approach to prove a generalization or that she automatically developed indirect reasoning because of it. Developing indirect reasoning does not only imply developing a conceptualization of proof as eliminating counterexamples, but also entails being able to provide mathematical reasons why the existence of counterexamples is indeed impossible, which involves further thinking. Moreover, it encompasses being aware of the characterization of all counterexamples. Concerning that aspect Yopp highlights as an important step that "to develop a proof using eliminating counterexamples approach, the arguer must view the approach as eliminating all possible counterexamples, which requires a general description of these objects." (p. 161). He also suggests that "constructing a description of possible counterexamples appears to be helpful in learning to reason with the 'impossible'" (p. 165), which is the case for the analysis of true universal statements and the impossible counterexamples for them. In my view, all this seems to explain the way Lizbeth came up with this idea to prove universal statements. The first two main discussions of the intervention were aimed at discussing false UASs and disproving them, which in our context inevitably involved explicit discussions related to the descriptions for all counterexamples that would refute the statements. In particular, Lizbeth was very active during those discussions (see Chapter 5, Section I.2.2.1), which might have supported her assumption dAL3[3] about proving universal statements.

Epp (2009) claims that "the more experience students have in seeing that a single counterexample disproves a universal statement, the more likely they are to understand

[255] Andrea did not use the term "semi-general counterexamples" (Peled & Zaslavsky, 1997), but the case she referred to qualified as one.
[256] In the introduction of Section I.2 of Chapter 5 I describe what I mean by "impossible counterexamples". In short, an impossible counterexample is a counterexample for a *true* universal statement.

that a general argument is needed to show that no counterexamples exists." (p. 314). Nevertheless, it is unclear whether Lizbeth saw a need for a general argument to show that no counterexample exists.

Another important case is Andrea's disproving of "some-statements", which involved her emergent development of an indirect form of reasoning. In order to disprove a "some-statement", Andrea focused on the impossibility that the "some-statement" was true (assumptions dAA7[6] and aAAt1). It shows again the importance of the connection between proving and disproving. The main influence of the intervention on her development of this assumption was her understanding of *what was sufficient to prove it* and ultimately *the logical interpretation of "some-statements"*. Andrea's original form of reasoning can be compared with Dubinsky et al.'s (1988) negation of the meaning method to negate a single-level quantification, but in the context of disproving and not of negating a "some-statement" (for details, see Chapter 5, Section II.2).

The minimal justification to disprove a US is a counterexample. During the intervention I tried to raise the teachers' awareness of this issue. The teachers also became aware of a *general characterization for counterexamples to UASs*. It allowed Andrea to move beyond and assume that a verbalized semi-general counterexample is valid to disprove a US (assumption aAAm1). During the intervention and before the episode where Andrea's assumption emerged we had not discussed any case of non-minimal justifications. Tsamir et al. (2009) report a case where a non-minimal justification was rejected as the justification did not follow the needed framework (it was more than a counterexample). In contrast, Andrea did not reject the non-minimal justification even though the given argument was different from the ones she was used to seeing when disproving USs. The form of expression in which the argument was presented only involved the use of text; however, Andrea could recognize that it was a counterexample for the given US. She called it a "verbalized counterexample" and accepted it as valid to disprove the statement in discussion (for details, see Chapter 5, Section I.2.2.2).

In this section I have shown three unexpected assumptions the teachers made during the 2018-intervention. Two of them are important as they suggest paths to begin to develop indirect forms of reasoning that, as has been reported, involve many challenges (e.g., Antonini & Mariotti, 2008). These assumptions show the importance of developing assumptions about what is involved in proving and disproving simultaneously. Likewise, Andrea's acceptance of a non-minimal justification for a false US reveals her attention to the general characteristics of counterexamples to the US and her insights about disproving USs. The three assumptions arose during the intervention without a pre-planned discussion to promote their emergence.

8. Summary of Section I

The conclusions in Section I are important as they point to issues that the mathematics education community should be aware of, related to assumptions that teachers use, as well as pointing to challenges that might arise as teachers engage in proof-related discussions. In such a context it is important to be aware that: teachers can make unusual assumptions, their assumptions might reveal connections they make between statements, their assumptions might exhibit that equivalent statements are treated differently, their mathematical assumptions might be based on non-mathematical reasons, their assumptions might involve the use of mathematical terms whose meanings are not mathematically aligned, their assumptions might be the result of overgeneralizations they make, and their assumptions might reveal the emergent development of further forms of reasoning.

It is important to be attentive to these cases as it can support a better development of the teachers' proof-related assumptions and not only at a local level as it has been usually addressed. With my research I have shown that teachers do not see statements in isolation. The teachers' assumptions show that developing assumptions about what is involved in proving and disproving ought to be taught together. They show the importance of aiming at consistency within the teachers' set of related assumptions. This should be considered when engaging teachers and others in proof-related discussions, where it is common to include specific discussions or tasks mostly (if not only) about universal statements (e.g., Chazan, 1993; Lee, 2016; G. J. Stylianides & A. J. Stylianides, 2009). Even if existential statements are included in a study (e.g., Barkai et al., 2002; Tabach et al., 2010b; Buchbinder & Zaslavsky, 2009), whether they initially see connections with universal statements or other forms of equivalent statements has not been investigated.

Future research on the development of teachers' proof-related assumptions and the design of future interventions for teachers, need to consider these assumptions, how they changed and their impact.

II. Conclusions about the features of the 2018-intervention that supported or hindered changes

In this section I focus on the features of the 2018-intervention in three cases: the features of the intervention that *supported changes* of the teachers' proof-related assumptions (Section 1), the features of the intervention that *hindered changes* of the teachers' assumptions (Section 2) and the features of the 2018-intervention that *fostered* the emergence of interesting *unexpected* proof-related assumptions (Section 3).

1. The features of the 2018-intervention that supported changes of the teachers' assumptions

The 2018-intervention aimed to support the development of the teachers' assumptions about mathematical modes of argumentation[257]. Not only that, but the main goal was that the teachers could explain why those mathematical modes of argumentation made sense, from a mathematical point of view; that is, the teachers should understand those modes of argumentation. My hypothesis was that the customized design for the 2018-intervention[258] would help me to achieve these goals.

In this section I focus on six features that contributed to the teachers' change of their non-mathematical assumptions: Drawing attention to the logical interpretation of SQ-statements (Section 1.1); challenging the teachers' assumption, which included: using cognitive conflicts, refining the concept "counterexample", returning a problem to the teachers, rejecting the teachers' non-mathematical assumptions (Section 1.2); using a variety of statements (Section 1.3); developing key proof-related concepts progressively (Section 1.4); expecting mathematical explanations to support their conclusions (Section 1.5); working with a small-group of teachers (Section 1.6).

[257] I use the expressions "modes of argumentation" and "forms of reasoning" interchangeably, as in A. J. Stylianides (2007).

[258] See Chapter 4, Section II.2.1, for details on the design of the intervention.

1.1. Drawing attention to the logical interpretation of SQ-statements

During the 2018-intervention I drew attention to the logical interpretation of single-quantified statements as a way to prompt the teachers to shape their own assumptions. Understanding the logical interpretation of SQ-statements implies noticing that the statement involves two sets: the set of analysis and the conclusion set; distinguishing the elements from the set of analysis to which the statement applies (i.e., whether the statement applies to *all* or *some* of those elements); and identifying the claim made in the statement about those elements (for details, see Chapter 3, Section II.1).

This feature is fundamental in the sense that challenging the teachers' non-mathematical assumptions (e.g., using of cognitive conflicts, Section II.1.2 below) relied on it. For example, to change Andrea's initial assumption dAA7[3] that if *"Some X are Y"* is true, then *"Some X are not Y"* is true, it was crucial that Andrea identified whether the statements involved in the conflicting situation were true or not. This first entailed her understanding of the logical interpretation of "some-statements" (for details, see Chapter 5, Section II.1).

Attention to their logical interpretation also supported proving "some-statements". Andrea's intervention-based assumption dAA1[10], that examples were insufficient to prove a statement, changed to her new assumption dAA7[4] that *"Some X are Y"* is true if there exists at least one confirming example for it. The new assumption was first introduced by Gessenia, whose explanation relied on the logical interpretation of the "some-statement" and convinced Andrea (for details, see Chapter 5, Section II.1).

Understanding the disproving of "some-statements" was also achieved by drawing attention to their logical interpretation. Andrea's initial assumption dAA1[1] that *"Some X are Y"* is false if *"All X are Y"* is true, first changed to another non-mathematical assumption. Andrea assumed that a "some-statement" was false if a "counterexample" could be found for it (assumption dAA7[5]). Even though she was hesitant about this assumption, it was clear that she was influenced by our previous discussions on universal statements and their disproving. Her assumption was rejected, which pushed her to look for another explanation. I drew the teachers' attention to the logical interpretation of the statement in discussion (*"Some X are not Y"* means that at least one *X* is not *Y*). This made Andrea consider whether it was possible for the statement to be true, which led her to conclude that *"Some X are not Y"* was false as it was impossible to find an *X* that was not *Y* (for details, see Chapter 5, Section II.2).

Andrea's and Gessenia's understanding of the sufficiency of one counterexample to disprove a UAS was supported by their focus on the logical interpretation of the UAS. Based on our discussions during the first part of the intervention the two teachers assumed that a counterexample was insufficient to disprove a US (assumptions dAA1[2] and dAG1[1]). The teachers recognized that a counterexample was sufficient evidence to disprove a UAS during the second part of the intervention as they paid attention to the logical interpretation of the UAS. Andrea explained why that mathematical evidence conclusively led to infer that the universal statement was false (for details, see Chapter 5, Section I.2.1).

The attention to the logical interpretation of a statement also guided Andrea's evolving understanding of when a negative "all-statement" is false. Andrea's assumption dAA7[11] that *"All X are not Y"* is true if there is an *X* that is not *Y*, followed from another assumption Andrea used, that *"All X are not Y"* is equivalent to *"Not all X are Y"* (assumption bAA8). As a result of the intervention, at that moment she already accepted that *"Not all X are Y"* is equivalent to *"Some X are not Y"*. Hence, she reasoned about the truth value of the negative "all-statement" through her analysis of the truth value of its equivalent negative "some-statement" (which was true because there was a confirming example for it). Andrea's assumption dAA7[11] changed as the teachers' attention was redirected towards the logical interpretation of the negative "all-statement" and what it meant for the statement to be true or false. Gessenia provided a counterexample and the teachers realized that the statement was indeed false, against their initial assumption. Andrea used that insight much later to disprove a similar negative "all-statement" (for details, see Chapter 5, Section III.2.2).

Particular attention to the "set of analysis"

The set of analysis was a construct I used during the 2018-intervention to draw the teachers' attention to the logical interpretation of SQ-statements. The *set of analysis* for a SQ-statement is the set to which the quantifier applies. For example, the set of analysis for the SQ-statement *"Some palindrome numbers are square numbers"* is the set of palindrome numbers (for details, see Chapter 3, Section II.1).

Attention to the set of analysis was particularly important for Gessenia and Lizbeth's change of their assumption bAG2[259] that a UAS and its converse state the same. Lizbeth's shift of assumption was inspired by Andrea's explanation, that if a UAS and its converse have different sets of analysis, or the conditions involved in those statements are different, then they cannot state the same thing. Andrea's argument exhibited her focus on the set of analysis to justify that a UAS and its converse did not make the same claim (for details, see Chapter 5, Section I.1.1.1). Gessenia's change was supported by her realization that the set of analysis of the converse statement may have elements different from those in the set of analysis of the original statement, which was against her initial assumption that both sets of analysis had the same elements (for details, see Chapter 5, Section I.1.1.2).

Attention to the set of analysis was also important to the teachers' refinement of the characterization for the counterexamples to USs, which includes Andrea's change of her initial assumption bAA4. Irrelevant examples (a type of challenger used) were rejected as counterexamples for the given universal statements as the teachers became aware that a counterexample should be an element that belonged to the set of analysis (for details, see Chapter 5, Section I.2.2).

Paying attention to the logical interpretation of a SQ-statement, and particularly to the set of analysis, was also useful in changing the teachers' initial assumptions bAA5 and bAL1 that confirming examples are sufficient to prove a universal statement. Tabach et al. (2010b) provide evidence of in-service secondary school teachers' awareness of the insufficiency of one confirming example to prove a universal statement. However, it is not clear where the teachers' awareness stems from and what its nature is. In contrast, with the teachers in my study I intended to show the way their assumptions evolved through our discussions and the possible different paths they took towards their understanding of the status of confirming examples when proving USs. For example, it is interesting that in contrast to my plans for the intervention, Lizbeth made her own decision and opted to use the term *justification* as *confirmation/verification* and to use

[259] Lizbeth exhibited the same initial assumption as Gessenia's; however, Lizbeth did not provide much details about its development. In those cases, I did not include a special section for the development of that teacher's assumptions in Chapter 5 and the teacher does not have a code for her assumption. Some details about Lizbeth's insights were included in the sections for the development of the other two teachers' assumptions.

other expressions (e.g., *guarantee*) to refer to the sufficiency of an argument to qualify as a mathematical proof. She then gradually adopted new meanings (for details, see Chapter 5, Section I.3.1.2). Finding out what individuals mean by the "justifications/arguments/proofs" they produce and whether they are aware of their role may not be trivial work (Hoyles & Küchemann, 2002; G. J. Stylianides & A. J. Stylianides, 2020; see Chapter 2, Section I.4). The particular case of Lizbeth shows that.

For Andrea, it was important to refine her assumptions related to proving USs by focusing on the set of analysis and the number of cases involved in them. Attention to this made her aware that confirming examples may be valid to prove finite USs with a "small" number of cases or to prove infinite USs when the examples are seen as generic. This allowed Andrea to make a distinction between finite and infinite USs and consider the status of confirming examples when proving them (for details, see Chapter 5, Section I.3.1.1).

In this section I have reviewed the evidence supporting my conclusion that drawing attention to the logical interpretation of SQ-statements was important in modifying the teacher's assumptions. The issue with using examples to prove USs is a well-established problem in the literature; however, I have shown that when attention is drawn to the logical interpretation of SQ-statements, teachers might focus on more refined issues like the *set of analysis* and whether the statement refers to *all* or *some* of those elements, as well as the number of cases involved in it. This shift in their attention helped them being more cautious about their decisions on what makes an argument a valid proof. Attention to the logical interpretation of statements also supported the teachers' realization that because the sets of analysis of a UAS and its converse are different, they cannot state the same thing and that one does not necessarily imply the other, which has been reported as a common persistent assumption that individuals use and for which G. J. Stylianides and A. J. Stylianides (2017) have noted that no meaningful advances have been taken at that time, and I am not aware of any since[260]. This feature was complemented with the progressive development of "some" feature (see Section II.1.4) to change Andrea' assumption that a "some-statement" cannot be universally true. They drew attention to the logical interpretation of a "some-statement" and what needs to be shown to prove it.

1.2. Challenging the teachers' assumptions

During the intervention I used different ways of challenging the teachers' assumptions, which had different characteristics and goals.

1.2.1. Changing assumptions through cognitive conflicts

I used the strategy of triggering cognitive conflicts on regular basis during the 2018-intervention. Whenever I noticed that the teachers' assumptions were contrary to standard mathematical assumptions, I tried to create cognitive conflicts[261].

Cognitive conflicts were particularly useful to change the following assumptions:

- The converse of a true universal statement is true (bAG1)
- The converse of a true universal statement is false (bAA2)

[260] See Chapter 2, Section I.1, for a review on the "converse error", possible explanations why this assumption might emerge and hints for possible solutions based on previous research.
[261] For details on the approach, see Chapter 3, Section I.1.1.

- Confirming examples are sufficient to justify that an infinite[262] UAS is true (bAA5 and bAL1)
- If *"Some X are Y"* is true, then *"Some X are not Y"* is true (dAA7[3])
- Confirming examples are insufficient to prove a statement (dAA1[10])

I challenged Gessenia's initial assumption bAG1 that a true universal statement implies that its converse is true, with the case of a true mathematical universal statement with a false converse; however, this was not enough to create a cognitive conflict in Gessenia. Three conditions were required for her change of assumption: (1) The background knowledge involved in the conflicting example had to agree with Gessenia's personal background knowledge for her to realize and understand the conflict; (2) she needed to grasp first the concept "set of analysis", which was supported with a *familiar* statement whose common interpretation agreed with its mathematical interpretation; (3) a conflicting familiar-context statement needed to be followed by an *imaginary* statement to support Gessenia's reflection on a related, and more fundamental, assumption that finally triggered the conflict. This implies that Gessenia had to analyze different types of universal statements before she realized, understood and could address the conflict (for details, see Chapter 5, Section I.1.1.2).

A cognitive conflict was also involved in Andrea's change of her initial assumption bAA2 that a true UAS implies that its converse is false. Her assumption was challenged by a true mathematical biconditional statement. As a result, Andrea reconsidered her initial assumption as she realized that there *may* be cases of true universal statements with true converses. Unlike Gessenia, Andrea's understanding of the concept "set of analysis" and her mathematical background knowledge were aligned with that required to grasp the conflicting example (for details, see Chapter 5, Section I.1.1.1).

Cognitive conflicts contributed as well to cases where the change of an assumption was a sequence of refinements, rather than the rejection of an initial assumption. For example, the teachers began the intervention with the well-known assumption that verifying examples was sufficient to infer the truth of a generalization (bAA5 and bAL1)[263]. A false universal conjecture that admitted confirming examples and whose truth value was not difficult to realize supported the teachers' emergent shift from their initial assumption. The "monstrous" statement *"For every natural number* n *(different from zero), when substituted in the formula* $1 + 1141n^2$*, the result is not a perfect square number"* (see G. J. Stylianides & A. J. Stylianides, 2009) was a conjecture the teachers formulated. It was included to reinforce their emergent assumption as it admits many confirming examples and can be disproved with a hard-to-find counterexample. These led the teachers away from their initial assumption to a more extreme assumption, that no statement could be proved with confirming examples. This new assumption dAA1[10] was in turn challenged with statements defined in finite sets, for which the exhaustion method could lead to a proof. The cognitive conflicts created by these statements led to a further refinement, the new assumption dAA3[1] that as long as the set of analysis is large, examples are not valid; if it is a small set, then they are. Arguments that included generic examples were discussed to further refine a new assumption dAA9[9]: When the statement is universal and true with infinite cases involved, an example is not a valid justification, unless it is a generic example[264]. Andrea was the only teacher who explicitly

[262] I use "infinite UAS" to refer to UASs that have infinite cases involved.
[263] See Chapter 2, Section I.2 for some prior literature where this issue is reported.
[264] In fact, we also discussed what might make a generic argument a proof. See Chapter 5, Section I.3.1.1 for details.

exhibited all these refinements (for details, see Chapter 5, Section I.3.1.1). Gessenia and Lizbeth understood that confirming examples may be insufficient to prove universal statements, but it is not clear if they understood the later refinements. Lizbeth in particular may have attended to other aspects at the time (for details, see Chapter 5, Sections I.3.1.2, I.3.2 and I.3.3).

I also used cognitive conflicts to support Andrea's refinement of her initial assumptions dAA1[8] and dAA7[3] related to "some-statements". Her assumption dAA1[8] revealed that her use of the quantifier "some" was consistent with the everyday language interpretation of "some" as "some, but not all" (e.g., Epp, 2003; Lee & Smith, 2009; Woodworth & Sells, 1935). She assumed that her everyday-life interpretation of "some-statements" could be applied to the mathematical context. The initial input I provided (in mathematics "some" means "at least one") did not conflict with her initial assumptions. An example of a true mathematical "some-statement" that was universally true triggered the cognitive conflict. It made Andrea modify her assumptions dAA7[3], that if *"Some X are Y"* is true, then *"Some X are not Y"* is necessarily true, and dAA1[8] (for details, see Chapter 5, Section II.1).

According to Posner et al. (1982), in order to achieve conceptual change through cognitive conflict, more than "metacognitive awareness" is needed. The individual needs to be aware that there is something s/he needs to change, but s/he also needs to be willing to do so. In the case of the three teachers who participated in my study, there was a high expectation for them to justify their assumptions. For a change to involve understanding, they needed to be able to explain why their new assumption made sense from a mathematical perspective (see Section I.1.5 below).

1.2.2. Refining the concept "counterexample"

This form of challenge was crucial to understanding disproving of universal affirmative statements and in particular to refining a characterization of counterexamples to a UAS. Andrea's initial assumption bAA4 that a counterexample is a conflicting example or situation, evolved when we discussed irrelevant and confirming examples. These two types of examples were expected to be rejected as counterexamples. In the process of explaining why those examples did not qualify as counterexamples, Andrea developed a general characterization for counterexamples to USs: A counterexample should satisfy the first condition of the statement, but contradict the second condition. Even though my target during the 2018-intervention was to tackle minimal justifications, Andrea's insight allowed her to go beyond and acknowledge the validity of non-minimal disproofs for universal statements (see Chapter 5, Section I.2.2.2 below).

Although the characterization of counterexamples to universal *negative* statements was not a focus of the 2018-intervention, Andrea managed to identify the characteristics that a general counterexample to a negative "for-every-statement" should have (what I called her "looking for counterexamples" approach in Chapter 5, Section I.2.2.2). Nevertheless, when the truth value of a negative "all-statement" was put forward, Andrea did not use a similar strategy. This suggests that either the form of the universal statement had an influence, or that other assumptions led her to conclude wrongly that the (false) negative "all-statement" was true were more predominant than a consideration that it could be false (see e.g., Chapter 5, Section III.2.2.2).

1.2.3. Returning a problem to the teachers

Giving back a problem to the teachers, similar to what Brousseau (2002) calls "devolution", was an approach I used to challenge the teachers' assumptions. With this strategy I attempted to promote discussions and create debates so that the teachers could refine their assumptions and find explanations for them on their own.

This approach supported Andrea's realization that a "some-statement" is proved to be true by showing the existence of at least one confirming example. This assumption is related to Andrea's change of her assumptions: *"Some X are Y"* and *"All X are Y"* both cannot have the same truth value (assumption dAA1[1]) and *"Some X are Y"* implies that *"Some X are not Y"* (assumption dAA7[3]), which are a direct consequence of her initial meaning for "some" as "some, but not all" (assumption dAA1[8]). Her intermediate assumption dAA1[10] in her development was that confirming examples were insufficient to prove a statement. Gessenia provided a confirming example to prove a "some-statement", which puzzled Andrea. Based on Andrea's reaction, I returned the problem to her, by asking the teachers whether that example was sufficient to prove the given "some-statement" ["*Is it (one confirming example) sufficient?"*] (for details, see Chapter 5, Section II.1).

1.2.4. Rejecting the teachers' assumptions

Rejections of the teachers' assumptions were also used during the intervention as challenges. For example, in the process of developing her assumptions for disproving "some-statements", Andrea hesitated about whether a "some-statement" could be disproved with counterexamples (assumptions dAA7[5a] and dAA7[5b]). My negative answer and remark that we had discussed the case of counterexamples for universal statements pushed her to think further in her search for other reasons to explain why a "some-statement" was false. This, plus her attention to the logical interpretation of the statement, led her to conclude that a "some-statement" was false if it was impossible to find a confirming example for it (for details, see Chapter 5, Section II.2).

Summary of Section II.1.2

In this section I have reviewed the evidence supporting my conclusion that challenging the teachers' assumptions was an important feature that supported changing the teacher's assumptions. This feature took different forms. Using cognitive conflicts supported the change of Andrea's assumptions related to the status of confirming examples when proving universal statements, Andrea's and Gessenia's assumptions about the relationship between a UAS and its converse, and Andrea's inferences with "some-statements".

It is important to underscore that the cognitive conflicts that I fostered during the intervention included conflicting examples that were enlightening for one teacher and not for the others. This seems to have been particularly the case for the conflicting example related to "some-statements" (for details, see Chapter 5, Section II.1). It is possible that it was a matter of time for Gessenia to realize the conflict as clearly as Andrea did. In addition, note that efforts to shift non-mathematical proof-related assumptions can create new inaccurate assumptions. Andrea's case for her transition from empirical arguments to her extreme rejection of confirming examples to prove universal statements is an example of this situation (see Chapter 5, Section I.3.1.1). It does not seem to have been taken into consideration in G. J. Stylianides and A. J. Stylianides' (2009) study. The instructional sequence described there may have created "intermediate" assumptions similar to those I found in my research. Andrea's case also suggests an extension of Buchbinder and Zaslavsky's (2009) framework for the status of confirming examples when proving universal statements. A refinement should be considered to include those cases where confirming examples may indeed prove universal statements.

Refining the concept "counterexample" promoted the teachers' modification of the initial given input for "counterexample" to refine a characterization for it, which was achieved with the discussion of irrelevant and confirming examples. Returning a problem to the teachers supported Andrea's reconsideration about the sufficiency of one confirming example to prove a "some-statement". Rejecting the teachers' assumptions pushed Andrea to search for different reasons for why a "some-statement" was false, which is a topic that is rarely addressed in school contexts (Tabach et al., 2010a).

1.3. The use of a variety of statements according to their content

Content has been identified as an important factor that influences the way individuals perform when solving logic-related tasks (e.g., Hawkins et al., 1984; Valiña & Martín, 2016). Even though the focus of the 2018-intervention was on mathematics, the statements I used during the proof-related discussions also included *familiar* and *imaginary* statements[265], each with specific goals.

These two types of statements were crucial for Gessenia's change of three of her assumptions. A familiar UAS supported Gessenia's realization that the UAS in discussion and its converse did not state the same (assumption bAG2) because the sets of analysis were different. Because of their generic nature, imaginary statements can focus attention on, for example, the general characteristics of a conflicting example in the context of a cognitive conflict. In Gessenia's case, an imaginary statement triggered a cognitive conflict with her assumption dAG2[2] that the sets X and Y involved in the statement *"All X are Y"* have the same elements, which was the basis for her assumptions bAG2 and bAG1 (the converse of a true UAS is true). Another familiar statement supported her change of assumption dAG2[2] that was also revealed through her change of assumptions bAG2 and bAG1 (see Chapter 5, Section I.1.1.2 for details).

Using statements with different content (mathematical, familiar, imaginary) was an important feature of the intervention. Notably, it supported Gessenia's change of her non-mathematical assumptions about the relation between a UAS and its converse. The idea of imaginary statements was inspired by Hawkins et al.'s (1984) *fantasy problems* (see Chapter 4, Section II.2.1; stage 1.3). Hawkins and colleagues found out that fantasy problems, when given first to preschoolers, elicited more logical responses. In contrast to one problem in which practical knowledge can have an influence, the content of fantasy problems did not interfere with their logical performance to solve syllogism problems. Although I see the relevance of Hawkins et al.'s result, I found out that in Gessenia's case it was important to analyze first a familiar statement that facilitates a particular observation and then an imaginary statement to introduce hesitation that can help her reconsider her initial non-mathematical assumption.

1.4. Progressive development of key proof-related concepts

In mathematics, the logical interpretation of a SQ-statement involves understanding some concepts. Notably, it entails understanding quantifiers from a mathematical perspective. An important feature of the 2018-intervention was to provide initial inputs for three

[265] For details on what I mean by *familiar* and *imaginary* statements, see Chapter 4, Section II.2.1, Stage 1.3.

important proof-related concepts: “proof” and “counterexample” and the quantifier “some”. The initial inputs were expressed in broad terms, and subsequent discussions aimed at supporting the teachers’ refinement of their conceptualizations.

The progressive development of proof-related concepts was important particularly for the teachers’ refinement of their concept “counterexample”. Andrea was the only teacher who began the intervention with an initial conceptualization for it. For Andrea, a counterexample was a conflicting example/situation (her initial assumption bAA4). Her personal usage of “counterexample” was not linked to the disproving of USs. Nonetheless, inspired by her initial conceptualization, I based the initial input on it (see Chapter 5, Section I.2.1) and the refinement of the concept was supported by discussions about irrelevant and confirming examples (see Section II.1.2.2 above).

Furthermore, the teachers’ paths for their conceptualizations of “proof” did not necessarily take the same course. For example, as they changed their initial assumptions bAA5 and bAL1 that confirming examples are sufficient to justify that an infinite UAS is true, Andrea’s focus was on the number of cases involved in the statement (see Chapter 5, Section I.3.1.1), whereas Lizbeth’s attention was directed towards the terms she used to distinguish validity (see Chapter 5, Section I.3.1.2) and the existence of counterexamples or not in order to prove a US (as I discuss further in Section I.3, below).

I gather and summarize the conceptualization of proof reached by the three teachers during the 2018-intervention in Figure 60. It is a composite conceptualization that integrates the three teachers’ personal conceptualizations. The three teachers did not necessarily finish the intervention with the same conceptualization.

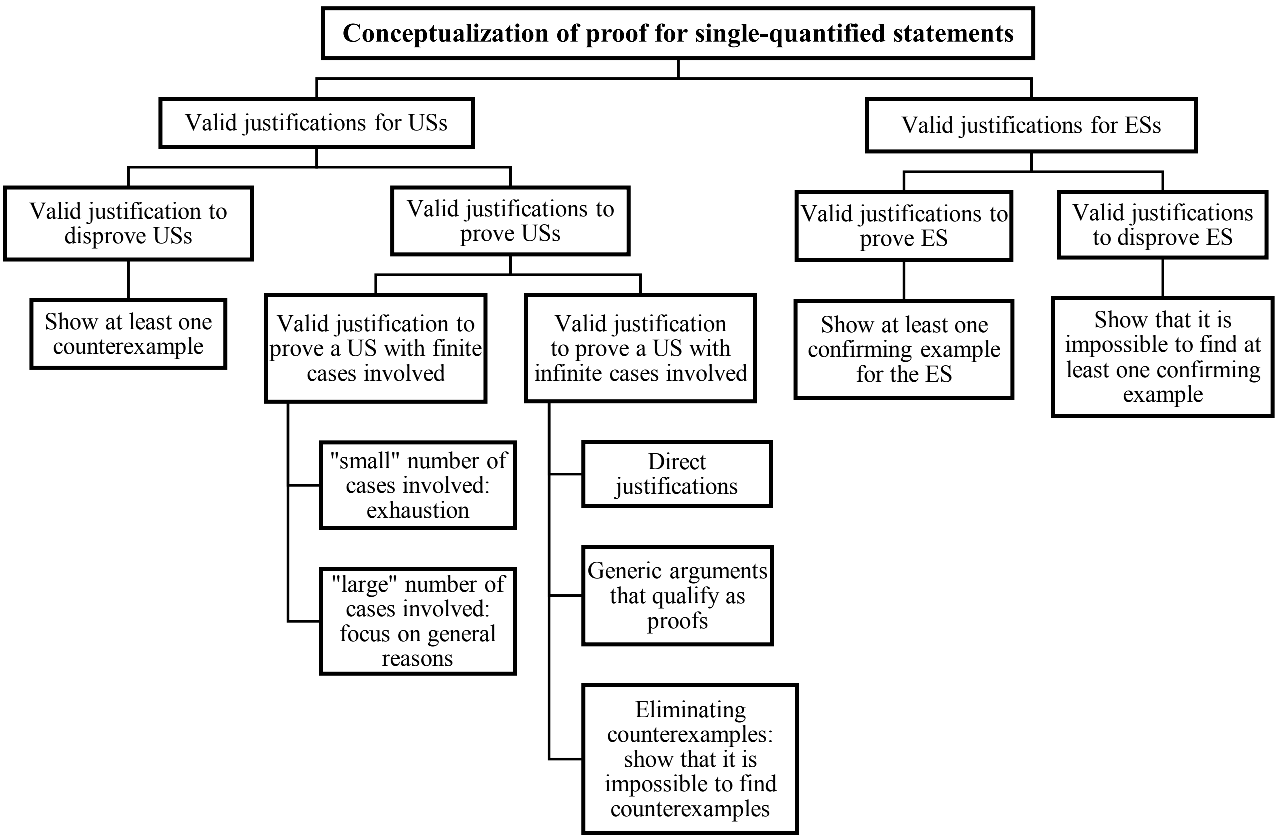


Figure 60. Conceptualization of proof developed by the teachers during the 2018-intervention

The progressive development of concepts also supported Andrea's change of her initial assumption dAA1[8] that "some" means "from all, one group, but not all", and the immediate inference she used for "some-statements"; namely, that *"Some X are Y"* implied that *"Some X are not Y* (assumption dAA7[3]). A cognitive conflict was embedded in this approach (see Section II.1.2.1 above), which triggered the shift.

As I have shown in Chapter 5, the initial inputs I introduced for the three terms were not sufficient for the teachers to change or shape their initial conceptualization. For example, this was the case with "some". The initial input I gave ("some" means "at least one") was not "totally" in conflict with the initial daily-life interpretation they used ("some" means "some, but not all"). This was especially true considering that the first examples we discussed of "some-statements" were true affirmative "some-statements" that were not universally true. In fact, other stimuli should support triggering those realizations. For instance, in the case of "some" and "proof", cognitive conflicts played a crucial complementary role, while in the case of "counterexample" challenging examples were determinant (see Section II.1.2.1 above).

This section shows that the progressive development of key proof-related concepts was an important feature of the 2018-intervention as it supported the teachers' emergent conceptualizations of "proof", "counterexample" and the quantifier "some". This feature was meaningful to the teachers because they had to make refinements on their own and explain them. With this approach they could progressively gain insights about these terms. They learned, for instance, that counterexamples to a UAS had specific characteristics, they were aware of them and made sense of why they described counterexamples to a UAS.

In fact, it is very difficult to describe the concept "proof" in a way that includes all its mathematically important aspects (Weber, 2014; Czocher & Weber, 2020). Based on my experience with the 2018-intervention, I believe that a conceptualization of "proof" ought to emerge gradually. This supported the teachers' refinement of their understanding of dis/proving. It allowed the teachers to make sense of when an argument is a valid dis/proof or not. They learned to make this distinction based on the logical interpretation of the SQ-statements and the conflicting cases we analyzed. The teachers developed their own conceptualization of "proof" and what is involved in dis/proving in a way that seems to be similar to a *cluster concept* or a *cluster category* (see Weber, 2014; Czocher & Weber, 2020). The properties of a justification that make it a proof do not need to be shared by all members of the category "proof". The conceptualization of proof is progressive as it is submitted to revisions and has a focus on increasing the understanding of the proof category.

1.5. The expectation for mathematical justifications to support their conclusions

The expectation for justifications is in fact a principle of proof-based teaching (PfBT)[266]. In the proof-based teaching I reported before (e.g., Vallejo-Vargas & Ordoñez-Montañez, 2014; Reid & Vallejo-Vargas, 2019) the focus was on mathematical facts (e.g., the sum of an even number of odd numbers is an even number) and their explanations. A common practice in PfBT contexts is the expectation for justifications of the conclusions that teachers and others reach. Those explanations are expected to be deductive, although there might not be an explicit introduction to deductive methods of proving. Explanations are continually refined in such a way that teachers and others can make sense of them and

[266] For some details on Proof-based Teaching, see Chapter 1 or Chapter 4, Section II.1.1.

fit the criteria of being deductive. With this feature, during the intervention the teachers became aware that their explanations needed to be mathematical in contrast to everyday forms of reasoning.

In the context of the 2018-intervention for teachers, the main focus was on establishing mathematical reasoning principles (assumptions) and explaining them from a mathematical point of view. The expectation for mathematical justifications contributed to make the teachers use the *mathematics register* (Halliday, 1978) as a frame. In particular, this was important to change Andrea's initial interpretation of the quantifier "some" as "from all, one group, but not all" (assumption dAA1[8]). Turning her attention to the mathematics register through the expectation for *mathematical* justifications pushed Andrea to use the initial input I provided for the mathematical meaning of "some", which supported her change of related assumptions (for details, see Chapter 5, Section II.1).

The expectation for justifications in the intervention drew the teachers' attention to constantly providing explanations. However, in certain contexts, the expectation for justifications might have been a tricky norm. For example, when evaluating others' arguments to disprove USs, Gessenia accepted *repetitive arguments*; that is, she accepted that the statement *"All X are Y"* is false, because not all *X* are *Y* (assumption dAG1[2]). Gessenia tried to introduce her awareness that the statement was false and why it was false to complete the given argument. This suggests that Gessenia might have assumed that to evaluate the argument, because of the expectation for justifications culture in the classroom, she was allowed to provide the missing explanations. Lizbeth's prompt supported Gessenia's awareness that those explanations were supposed to be given by the author of the argument, not by her (for details, see Chapter 5, Section I.2.1.1). In Morris' (2007) terms, Gessenia had to "monitor the introduction of personal knowledge" (p. 484). She had to learn to limit herself to evaluate arguments as they were given.

In this section I have recalled the evidence that showed why the expectation for mathematical justifications was an important for the 2018-intervention. This feature promoted the teachers' shift of their attention to the mathematics register. It was complemented in several cases with the introduction of initial inputs for the meaning of key terms as starting mathematical points from where to begin to reason (see Section II.1.4). Learning mathematics, and notably about proof, requires learning its language (Schleppegrell, 2007). I have shown that this feature contributed, for example, to Andrea's change of her assumptions about "some-statements" given that Andrea had to rely on the *mathematical* meaning of the quantifier "some".

1.6. The small-group context

The intervention took place in a small-group context, in which the teachers participated by sharing ideas and supporting each other while forming their assumptions. The group was the right size to encourage the teachers to speak up, ask questions, listen to others' ideas and be reflective. Furthermore, the teachers did not feel "evaluated", as they might have felt if they were in an interview-like environment. In particular, the small group context supported the teachers' insights as they had several opportunities to learn from each other. For example, this was relevant for the reconsideration of Gessenia's assumption dAG1[2] that a repetitive argument can disprove a universal statement. The change was successfully supported by Lizbeth, who suggested to Gessenia that she should not complete the given argument by introducing her personal knowledge when evaluating it.

The small-group context also supported Andrea's change of her assumption dAA7[11] that *"All X are not Y"* is true because it has a confirming case (an *X* that is not *Y*). Gessenia also shared the same initial assumption, even though the reasons she used to make sense of the assumption were probably not the same as Andrea's. Gessenia's active participation during the discussion was crucial to shifting this assumption. During the intervention Gessenia realized that the statement being considered was false, as it had a counterexample. In that process she first provided confirming examples for it and seemed convinced that the statement was true. Gessenia's process of realization persuaded Andrea to change her initial stance (see Chapter 5, Section III.2.2.2).

For Lizbeth and Gessenia it was also beneficial to interact in a small group with Andrea. Andrea was able to formulate a description of counterexamples to UASs in general terms (see Section I.1.2.2 above in relation to Andrea's assumption bAA4), to which Lizbeth and Gessenia were possibly not paying attention (see Chapter 5, Section I.2.2.2).

Gessenia's change of her assumption dAG1[4] that a counterexample should not satisfy at least one condition of the statement was also supported by the small-group context. Observing Andrea and Lizbeth's gaining insights about the characterization of counterexamples supported Gessenia's shift (see Chapter 5, Section I.2.2.3).

It may not be a surprise that working with a small group of teachers was beneficial for the development of the teachers' assumptions as it was almost a personalized interaction. However, this feature has shown to be beneficial during the 2018-intervention when there was a leading voice that persuaded others. In some cases this depended on the topic of discussion. Yet, if there had not been a teacher who made her claims explicitly and provided explanations for them (because of the expectation for justifications feature, see Section II.1.5), it is likely that this feature would not have been as productive as it was.

The small-group context was an important feature of the 2018-intervention as it promoted changes of Gessenia's assumption related to the sufficient mathematical evidence to disprove a US, Andrea's assumption about the truth value of a negative "all-statement", and Lizbeth's and Gessenia's assumptions about the characterization of counterexamples to UASs. This feature was supported by discussions in which the development of the teachers' assumptions was explicit. The discussions allowed the teachers to pay close attention to the way other teachers developed their own reasoning and gain insights from them. I argue that the small-group context was beneficial for two reasons: (1) the teachers did not feel intimidated due to a big number of participants, and (2) observing/listening to their colleagues discuss mathematical issues and progressively develop their reasoning showed to them that it was a real process of beginning to understand mathematics, instead of being told what to know, which is common in vertical learning environments, where the teacher is the provider of knowledge.

1.7. Summary of Section II.1

Section II.1 highlights the features of the 2018-intervention that supported changes of the teachers' non-mathematical assumptions. The intervention drew the teachers' attention to the logical interpretation of SQ-statements, challenged the teachers' assumptions in different forms that included the use of cognitive conflicts, engaged the teachers in discussions about statements with different content, promoted the gradual development of key-proof related concepts, reminded the teachers of the expectation for mathematical justifications to support their claims and involved a small number of teachers, which facilitated our discussions and forming their insights.

These features did not provoke changes on their own. They worked together to achieve that goal. For example, as I mentioned before, drawing the teachers' attention to the logical interpretation of statements was a fundamental feature as other features depended on it. Notably, it contributed to provoking a cognitive conflict and change Andrea's assumption related to inferences she had assumed as obvious between "some-statements". In addition, the cognitive conflict was supported by the gradual development of the concept "some", which is another feature of the intervention. All features were strengthened by the expectation for justifications created in the classroom. This feature was fostered because of the proof-based teaching context in which the intervention was developed and favored the teachers' reflection on their claims. These features are all relevant and clearly interconnected.

The mathematics education community needs to be aware of these features and explore their implications for future interventions with other groups of teachers.

2. The features of the 2018-intervention that hindered changes of the teachers' assumptions

The 2018-intervention also had features that impeded that the teachers could change their non-mathematical assumptions. In this section I focus on four features: Introduction of negation as a rule (Section 2.1); insufficient attention to language (Section 2.2); insufficient emphasis on the various forms in which statements can be presented and the connections between them (Section 2.3); lack of awareness of fundamental assumptions and their impact (Section 2.4). All these features are related to two topics: negation and "no-statements".

2.1. Introduction of negation as a rule

Introducing rules may make it difficult promoting understanding. Further, as I showed in Chapter 5 and recall here, it may create additional conflicts.

My study showed that Andrea held some initial assumptions for some simple implicit negations, from which other related assumptions arose. For example, the case of *"Not all X are Y"* exhibited Andrea's assumption that:

- *"Not all X are Y"* is the same as *"Some X are Y"* (dAA7[1]),

which was directly linked to her initial assumptions related to "some-statements":

- If *"All X are Y"* is true, then *"Some X are Y"* is false (dAA1[1])
- If *"Some X are Y"* is true, then *"Some X are not Y"* is true (dAA7[3])
- "Some" means "from all, one group, but not all" (dAA1[8]).

The 2018-intervention had the goal of introducing existential statements as the simple implicit negation of universal statements. I did not expect to go into many details about negations; however, the teachers' questions, requests and doubts led us to engage in several related discussions. Those discussions revealed that Andrea already carried her own initial assumptions about negations, and also made me aware that some of them were linked to language (see Section II.2.2 below), which directly influenced the way she negated.

When I became aware of Andrea's initial negation of USs (assumption dAA7[1]), I tried to change it. I introduced the following rule and expected the teachers to reflect on it afterwards:

> The negation of a universal statement results in an existential statement and the negator also negates the consequent.

According to this rule, the simple implicit negation *"Not all X are Y"* was the same as *"Some X are not Y"*, or equivalent existential statements, which differed from Andrea's negation dAA7[1]. In fact, Andrea used an initial general approach to find equivalent statements for simple implicit negations. I called it her "Separate and Substitute" (SS-) approach. Recall that it consists of separating, for example, *"Not all X are Y"* as *"[Not all X] are Y"* and substituting "not all" with "some" (her *semantic substitution* for "not all", see Dawkins & Cook, 2017), which resulted in the statement *"Some X are Y"*[267].

The rule did not immediately change Andrea's initial assumption dAA7[1] because of her non-mathematical assumption dAA7[3]. Andrea kept her initial assumption dAA7[1] until her assumption dAA7[3] and dAA1[8] changed. When they changed, Andrea's initial assumption dAA7[1] shifted; however, its change did not involve any understanding as she merely adapted her initial SS-approach to obtain the same negation as the rule. I call this modification Andrea's DSS-approach, where DSS stands for Distribute, Separate and Substitute. She used her new approach to obtain expected mathematical negations, although it was only a shortcut that still involved her use of a non-mathematical semantic substitution of "not all" with "some".

Another important case was the negation of existential statements. In hindsight, it is likely that Andrea's lack of understanding of the negation of ESs was due to its introduction as a rule, plus two other factors: (1) the rule I provided was vague [*the negation of an ES is going to affect both, antecedent and consequent*]; (2) Andrea's lack of a fundamental understanding of the logical interpretation of "no-statements". The rule to negate ESs did not specify the way the antecedent and consequent were affected. In contrast, the rule I gave for USs explicitly indicated that the negation of a US was an ES and that its consequent had to be negated.

Andrea's lack of understanding of the negation of existential statements revealed her current non-mathematical assumptions about "no-statements". In particular, Andrea did not hold initial assumptions for the negation of "some-statements". With the intervention Andrea *accepted* that the negation of statements of the form *"Some X are Y"* was *"No X is Y"*; however, the evidence suggests that she merely treated it as a rule that did not involve any understanding (for details, see Chapter 5, Section II.3.1).

Furthermore, introducing negations as rules turned out to be especially conflicting in two specific cases: the negation of "no-statements" and the negation of negative "there-exist-statements". Andrea's initial assumptions dAA9[6] and dAA7[22] were mathematically aligned; that is:

- *"There does not exist X that is not Y"* is the same as *"No X is not Y"* (dAA9[6])
- The negation of *"No X is Y"* is *"Some X are Y"* (dAA7[22])

[267] Andrea also used her SS-approach for the simple implicit negation of "there-exist-statements" and for the explicit negation of "no-statements", for which she had available semantic substitutions. The semantic substitutions she used in those cases were "there does not exist" as "no", and the negation of "none" is "some" (for details see Chapter 5, Sections II.3 and III.1.3).

Yet, her overextension of her DSS-approach to these cases led her to obtain a different equivalent statement for *"There does not exist X that is Y"* and a different negation for *"No X is Y"*; namely:

- *"There does not exist X that is not Y"* is the same as *"No X is Y"*
- The negation of *"No X is Y"* is *"Some X are not Y"*

To Andrea this clearly felt as a conflict that puzzled her and that was not resolved during the intervention.

The importance of negations in mathematics has been widely recognized (e.g., Bardelle, 2013; Dubinsky et al., 1988; Epp, 2003; Sellers, 2020)[268]. However, individuals usually rely on non-mathematical procedures to negate (e.g., Barnard, 1995; Pasztor & Alacaci, 2005). In particular, Barnard (1995) stated that before the participants were requested to choose negations, they were reminded the meaning of "negation" and received a list of examples; however, that information was not provided in their report. In contrast, Pasztor and Alacaci (2005) included the way negation was prompted in their study: "*negation* of a sentence A was defined as a contradictory sentence B that denies the truth of the given sentence A, and in every situation, one of A and B must be true and the other false" (p. 1716). Although teachers and others might need some starting point to begin to understand negation, Pasztor and Alacaci's prompt might be interpreted in different ways according to, for example, the interpretation the reader gives to the term "contradictory".

Dubinsky et al. (1988) has already observed that negating by rules is a mechanical method that relies on retrieving rules from memory, and can easily fail. Not only that, if understanding negation is intended, introducing negations as rules should be avoided. It might not only discourage the search for understanding, but can also introduce new conflicts as I have shown with my research. During the 2018-intervention Andrea found no reason to make efforts to understand negation. The mechanism she used to obtain negations that matched my expectations was making a rule of her own, which was based on her initial non-mathematical approach and the rules I gave. Although in the end Andrea managed to obtain mathematical negations of "all-statements" and "some-statements" with her (DSS-) rule to negate, her negation did not involve understanding. Moreover, her negation of "there-exist-statements" introduced conflicts.

2.2. Insufficient attention to language

Language is used as a means to communicate with each other. Sometimes communication is hindered because of individuals' different uses of language. There were some language-related issues that I did not take into consideration during the intervention and which could have had a positive effect in the development of the teachers' understandings if I had done so. Here I give examples of proof-related terms Andrea used in their daily-life and that she carried those meanings to our mathematical discussions.

The way some terms or expressions are used in everyday language can influence the way those expressions are used in mathematics (Halliday, 1978). A clear initial assumption Andrea used was her distinction between the quantifiers "all", "some" and "none", which was grounded on her everyday use of them.

- "Some" means "from all, one group, but not all" (dAA1[8])

[268] See Chapter 2, Section I.8, for details on previous research about Negation.

- *"No X..."* refers to everything that is not *X* (dAA7[12])

Andrea applied both initial assumptions (dAA1[8] and dAA7[12]) to her interpretation of "some-statements" and "no-statements" in a mathematical context. This influenced her assumptions related to "some-statements" as I have shown before (see assumptions dAA1[1], dAA7[1] and dAA7[3]) and those about equivalences with "no-statements" and their representations[269]; namely:

- *"No X is Y"* is different from *"All X are not Y"* (dAA7[20])
- The representation of *"No X is Y"* and *"No X is not Y"* is the same, disjoint sets *X* and *Y* (dAA7[13] and dAA7[25], respectively)

In addition, the teachers tended to rely on the verification of one sub-set of the infinite cases involved in a statement when asked to prove it (assumptions bAA5 and bAL1). This behavior may be explained in terms of the way the term "proof" is used in other fields, such as chemistry or physics, which might have been closer to their personal background knowledge[270]. Considering that the teachers were novices at *mathematical proof*, they had no other reference framework to depend on than the use of "proof" that was available to them.

Andrea's disproving of the implicit negation of a negative "there-exist-statement" seems to have been influenced by the way daily-life language is used. Andrea initially assumed that *"There does not exist X that is not Y"* was false because there exists an *X* that is *Y* (assumption dAA9[5]). This means that Andrea disproved *"There does not exist X that is not Y"* as if she was disproving *"There does not exist X that is Y"*. In everyday conversations in Spanish there is a phenomenon that is used to emphasize a negation: the reinforcement of a negation. It consists of using a second negation particle in order to reinforce a negation, which certainly may have an effect on the way mathematics statements were treated[271]. This might explain why Andrea disproved the statement *"There does not exist X that is Y"* when she had to disprove *"There does not exist X that is not Y"*. She might have assumed that the second negation particle was used to emphasize the first negation and as such it could be ignored.

Another factor that might have influenced Andrea's initial assumptions about negations is that some negation particles in ordinary language can change their place in the sentence without altering its meaning. For example, the sentences "It is not true that I am worried" and "I am not worried" have the same interpretation. This might have played a role in Andrea's initial assumption bAA8 that *"All X are not Y"* is the same as *"Not all X are Y"* (for details, see Chapter 5, Section III.2.2).

In this section I have shown that three aspects related to language had several implications in the teachers' proof-related assumptions during the 2018-intervention: the interpretation of quantifiers in everyday language, the way negation is emphasized in common language and the irrelevance of where the negation particle is placed when negating in common language.

Some of these issues have been already reported in previous research as problematic. The interpretation of the quantifier "some" might influence assumptions related to the disproving of "some-statements" (e.g., Tall, 1977). Students assume that negating a

---

[269] For details, see Chapter 5, Section III.1.1.

[270] For details on this issue, see Chapter 5, Section I.3.1.2.

[271] For details on Andrea's disproving of simple implicit negations of "there-exist-statements" and examples of the reinforcement of a negation, see Chapter 5, Section II.3.2.

statement involves negating only one part of the statement (e.g., Barnard, 1995; Dawkins, 2017; Dubinsky et al., 1988; Lin et al., 2003; Pasztor & Alacaci, 2005; Sellers, 2018)[272].

Lee and Smith (2009) have suggested developing research to "investigate the relation between students' interpretations and proof constructions, and also to see how instructional practices can attend to the cognitive challenge faced by students who are new to 'the rules of the proving game'" (p. 25). However, I am not aware of research that has taken a step forward to address these issues. Notably, little (if any) attention has been given to the way teachers interpret the quantifier "none" in everyday language and its implications in teachers and others' proof-related assumptions, or how the everyday interpretation of "some" influences different aspects about "some-statements" and their connection.

During the 2018-intervention, my insufficient attention to these aspects did not support the teachers' change of related non-mathematical assumptions.

2.3. Insufficient emphasis on the various forms in which equivalent statements can be presented and the connections between them

This feature refers to the equivalent forms in which statements can be presented. For example, a UAS can be expressed as *"All X are Y"*, *"Every X is Y"*, *"If X, then Y", "Each X is Y"*. Likewise, ENSs can be presented in many equivalent forms, such as *"There exist X that is not Y"*, *"Some X are not Y"*, *"At least one X is not Y"* and *"There is X that is not Y"* (for more details, see Chapter 3, Section II.2).

Being aware of the different equivalent forms a statement can take and using those equivalences to foster understanding may have played a role during the intervention. For example, this could have been useful in particular to the understanding of disproving UNSs and ESs.

During the intervention Andrea spontaneously used her "search for counterexamples" approach to find out whether a UNS of the form *"For every x in M, G(x) is K"* was false, even though we had only addressed falsity for UASs up to that moment. Through this process Andrea revealed her conscious characterization of counterexamples for a negative "for-every-statement". Nevertheless, when a UNS of the form *"All X are not Y"* (a negative "all-statement") was discussed much later, she did not even consider the possibility for it to be false (which it was). Andrea first used her initial assumption bAA8 that *"All X are not Y"* was equivalent to *"Not all X are Y"* and analyzed the latter instead, which took her to incorrectly conclude that the statement was true (see Chapter 5, Section III.2.2).

Implicit negations can look more or less simple depending on the form of the statement to be negated. For example, the implicit negation of "there-exist-statements" has a simpler form than that for "some-statements". *"There does not exist X that is Y"* is simpler than *"It is not the case that some X are Y"*. Further, in Spanish the former looks much simpler than the latter. This implies that disproving the simple implicit negation of "there-exist-statements" may be easier than disproving the negation of "some-statements". That was the case of the simple implicit negation of affirmative "there-exist-statements", for which Andrea assumed that:

- *"There does not exist X that is Y"* is false if there exists an *X* that is *Y* (dAA9[2])

[272] For details on this problem, see Chapter 2, Section I.8.

Taking advantage of the equivalences between ESs and the simplicity that "there-exist-statements" provide, understanding the negation of "some-statements" and its disproving could have been simplified. Nevertheless, Andrea interpreted "there-exist-" and "some-" statements differently (see Section I.3) and so the simple form of "there-exist-statements" could not support further understandings of "some-statements". There was a need to establish that linkage before, which was not accomplished during the intervention.

This section shows that the insufficient attention to the connections between equivalent statements in different forms hindered changes of the teachers' non-mathematical assumptions. In Section I, I have shown that the teachers make connections between statements. In Section II I have evidenced that they might treat equivalent statements as if they were not. Explicit connections between equivalent statements needs to be done. Andrea's disproving of negative "for-every-statements" might have supported her change of (or cast doubt on) the initial assumptions she used to determine the truth value of a negative "all-statement" if explicit links between the two UNSs had been previously established during the intervention. The simple form that the implicit negation of "there-exist-statements" might have simplified Andrea's negation of "some-statements" if she had seen the connections between the two existential statements.

This issue has received little (if any) attention in the research literature related to proof. Epp (2009b) has drawn attention to "translating back and forth from formal mathematical statements to their many different informal versions" as a way to support students' production and evaluation of proofs. Epp highlights the value of this activity for the case of definitions as she considers that the "truth or falsity of a mathematical statement is more apparent if one uses one phrasing of a definition rather than another" (p. 316), but her observations also apply to the case of SQ-statements. Selden and Selden (1995) have explored the way undergraduates unpack informal to formal calculus statements. They found out that no student could consistently do so even though they were familiar with symbols of predicate calculus.

As I have shown in my work, teachers and possibly others might not see connections even between equivalent single-quantified informal statements. There is no study that came to my attention that has had a focus on addressing this lack of attention as a way to facilitate the transfer of understanding of one form of statements to equivalent ones and support the change of teachers' non-mathematical assumptions.

2.4. Lack of awareness of fundamental assumptions and their impact

In some cases, I was not aware of the assumptions that were the basis of other assumptions and their impact, which did not allow me to address them quickly enough.

As I have shown in Section I.4, mathematical assumptions do not necessarily mean mathematically grounded assumptions. It is important to find out the nature of not only non-mathematical assumptions, but also the origins of mathematical assumptions to avoid that other assumptions be based on possible non-mathematical grounds. For example, Andrea's initial assumption dAA7[13] referred to the representation for the statement *"No X is Y"* and was, as expected, two disjoint sets *X* and *Y*. Nevertheless, her initial interpretation of "no-statements" on which she based her representation, was non-mathematical (dAA7[12]: *"No X..."* refers to everything that is not *X*). At that moment I was not completely aware of the impact that her initial resistant-to-change interpretation of "no-statements" would have and given that she provided a correct representation, I continued with the discussions. This, however, made subsequent related discussions more complicated (see Chapter 5, Section III.1.1).

There were some initial assumptions that were only accessible late in the intervention and were needed to "connect the dots". For example, Andrea's initial interpretation of the simple negation "not X" (assumption aAAm3) completed my understanding of the way Andrea reasoned about and represented "no-statements"; however, assumption aAAm3 was explored and exhibited only after the intervention, so I was not able to address it when assumptions about "no-statements" were analyzed during the intervention.

- Given two disjoint sets *X* and *Y*, *Y* represents the set of the elements that are "not *X*" (aAAm3)

Paying insufficient attention to what the teachers focused on is a tricky feature that sometimes is clearer when analyzed in retrospect. For instance, Andrea's representations of "no-statements" of the forms *"No X is Y"* and *"No X is not Y"* were identical (two disjoint sets *X* and *Y*). The reasons behind them were not evident to me at the time our discussions took place. Only in hindsight could I make sense of them.

In this section I have reviewed the evidence that shows that a lack of awareness of the teachers' fundamental assumptions might impede properly addressing the root obstacles to change other related assumptions. As shown in Section I.2, teachers make connections between statements involved in proof-related assumptions. Their assumptions usually rely on the logical interpretation of the statements involved and ultimately on the interpretation of the quantifier in use. If those basic assumptions are not tackled during an intervention, it is possible that other assumptions that derive from them might not change. A clear example is Andrea's assumptions about "no-statements" and the challenges that were involved when attempting to change her non-mathematical assumptions about them. During the 2018-intervention I was not aware of the implications that Andrea's initial use of the quantifier "no/none" would have for her set of related assumptions, especially considering that she used several mathematically aligned assumptions about "no-statements". This impeded me from addressing this issue properly.

2.5. Summary of Section II.2

Section II.2 highlights the features of the 2018-intervention that hindered changes of the teachers' non-mathematical assumptions or mathematical assumptions with non-mathematical reasons. During the intervention, negation was introduced as a rule, language did not receive enough attention, neither the various forms in which statements can be presented, nor the connections between them were emphasized enough, there was a lack of awareness of some fundamental assumptions and their impact.

The features in this Section II.2 are mainly related to Andrea's[273] assumptions about negations and "no-statements". A lack of awareness of Andrea's fundamental assumptions about them did not allow me to properly address them. On the one hand, her fundamental assumption about "no-statements" was only accessible after the intervention much later than her related assumptions. On the other hand, her initial semantic substitution approach to negate involved the use of semantic substitutions that were directly linked with her initial interpretation of quantifiers. Introducing negations as rules

[273] Andrea's assumptions were not representative of the three teachers' assumptions. This section is mainly about her since she engaged actively in the discussions and exhibited her assumptions explicitly.

stole any possibility of motivation that Andrea could have had to make sense of it. In both cases, my lack of attention to language (her initial interpretation of the quantifier "no" and her semantic substitutions) made it more difficult to be able to change Andrea's related assumptions. Because Andrea did not change these, she did not change either her non-mathematical related assumptions. The insufficient attention to language also impeded addressing Andrea's assumption about negative "all-statements" for which she seems to have found irrelevant where the negation particle was located. The little emphasis on establishing connections between statements in alternative forms did not allow me taking advantage of the statements whose forms of expression were easier for Andrea as revealed through her mathematically aligned assumptions.

III. Conclusions related to the ways the teachers' assumptions that changed were visible during their teaching

Research question RQ2 has a focus on the assumptions the teachers revealed during their teaching.

> RQ2: How are the in-service primary school teachers' assumptions that changed visible during their teaching in schools?

The proof-related activities I proposed for the teachers' teaching included identifying patterns, formulating conjectures, establishing the truth value of statements, proving and disproving. The statements involved in the activities were both universal and existential. In contrast to the intervention, the teachers' lessons did not include discussions about *negation*, which means that the teachers' teaching did not reveal the effects of the intervention in that respect.

In this section I focus on four methodological approaches that the teachers adopted from the intervention and applied to their teaching; namely:

1) Suggesting irrelevant and confirming examples can support the refinement of the conditions that counterexamples should or should not satisfy.
2) Drawing their students' attention to the logical interpretation of statements is vital to dis/proving.
3) Emphasizing the number of cases involved in a statement to support the identification of the sufficient evidence to prove a US.
4) Emphasizing the need to provide evidence when disproving USs.

The teachers' use of these methodological approaches reveal not only their proof-related assumptions at that time, but also the important role they attributed to the approaches and how the intervention changed their teaching. The teachers' assumptions that were observed during their teaching are in harmony with those they developed during the intervention.

1. Their use of confirming and irrelevant examples to support the refinement of the conditions that counterexamples should and should not satisfy

The use of confirming and irrelevant examples is a methodological approach that the teachers interiorized from the intervention and was common in all three cases. The three teachers used the same approach I used with them during the intervention to refine a general description of counterexamples. The teachers generated irrelevant and confirming examples on the spur of the moment, which they suggested to their pupils with the aim at being rejected as counterexamples that disproved the USs in discussion. Even though this was not planned as a teaching approach for the teachers' teaching, it was interesting that the teachers integrated it into their own class discussions. This suggests that it was a meaningful experience the teachers had during the intervention as they spontaneously replicated it in their own classrooms. I believe that they adopted the approach from the intervention since they found that it was not difficult to grasp and supported their goals. The fact that the teachers used this approach during their own teaching is evidence of its relevance when determining whether certain examples disprove or not certain universal statements.

The teachers' assumptions about the characterization for counterexamples to UASs were exhibited through their own teaching. In particular, the following assumptions were exhibited through the teachers' teaching:

- A confirming example cannot be a counterexample to a UAS because it supports the statement (dAL1[2])
- An irrelevant example cannot be a counterexample to a UAS because it does not satisfy the first condition of the statement (dAL1[4])
- A counterexample must satisfy the first condition, but contradict the second condition of the universal statement (dAA1[9] and aAGt2)

To support their students' understandings the teachers resorted to the logical interpretation of the discussed statements in order to underscore the conditions that possible counterexamples should have, which is another methodological approach they adopted from the intervention (see next).

2. Their focus on the logical interpretation of statements to develop understanding of dis/proving

A second methodological approach that the teachers gained during the intervention was the importance of understanding the logical interpretation of the statements in discussion (see Section II.1.1 above). The teachers relied on it, which included their attention to the set of analysis and the quantifiers involved[274]. For example, if the teachers attempted to engage their students in a discussion about a "some-statement", they contrasted it with an "all-statement". This means that they emphasized the difference between universal and existential quantifiers to support their students' understanding of USs, ESs and their dis/proving. Notably, Andrea made this distinction during her teaching that had a focus on the existential statement *"Some divisions by 4 have a remainder equal to 7"*.

> Andrea: *Is there indeed a division like that? When I say "some", I mean "at least one". When I divide by 4, is the remainder going to be equal to 7? Let's see, why?... Here, were you asked for ALL divisions?... It means, if there was at least one, if one holds, Giussepe [calling for one student's attention], if one holds, this is valid. At least one division, it says, one division by 4, that has a remainder equal to 7. Let's see whether this is true or not… there will be some divisions, here it says, such that when you divide by 4, the remainder will be equal to 7? There will be some?*

[274] For details on what I mean by understanding the logical interpretation of single-quantified statements, see Chapter 3, Section II.1.

Andrea's explanation made visible for example two of her new assumptions: one confirming example is sufficient to prove a "some-statement" (dAA7[4]), and *"Some X are Y"* is false because it is impossible to find an example of *X* that is *Y* (aAAt1) (for details, see Chapter 5, Section II.1).

The teachers' focus on their students' mathematical interpretation of statements was particularly useful to explain why one counterexample is sufficient to disprove a US. For instance, Gessenia relied on the logical interpretation of the UAS *"All natural numbers are divisible by 4"* to explain her class why one counterexample was enough to disprove it.

> Gessenia: *It says ALL. Look what it says here, ALL natural numbers. Daniel, I have one question, how many natural numbers are there? ... The natural numbers are infinite. Here it says ALL of them, ALL those infinite numbers, it says, are divisible by 4, and you all have said that this is not true. Here he [one student] gave [she refers to number 21], what? How have we called this? ... A counterexample... He says, not, and how many examples did he need to show this? ... One counterexample, the contrary. He is proving with a counterexample, just one. Miss, but there are many others. One-, one counterexample that breaks what is stated here is sufficient, ok?*

Gessenia's explanation exhibited for example the use of her current assumption about the sufficiency of one counterexample to disprove a US: when it says ALL and it is not true, it is sufficient to show a counterexample, and that there might be many counterexamples, but one is sufficient to disprove an "ALL-statement" (aAGt3, for details, see Chapter 5, Section I.2.1.1).

Andrea explicitly resorted to the "set of analysis" and the quantifier involved to explain why a statement was universal and when it was false. The "set of analysis" concept was introduced during the intervention for teachers and its explicit usage during Andrea's teaching reveals the relevance that she attributed to it. Her teaching exhibited her focus on the "set of analysis" of the statement *"All natural numbers are divisible by 4"* (St143) to disprove it.

> Andrea: *Kids, something he [student C] was telling us before and that you should always remember, when I have a statement like this one, and it is universal, universal because it begins with "all", because your set of analysis, what you are supposed to analyze are AAAAALL the natural numbers, when you have a set of analysis that refers to all, then that is a universal statement. When your universal statement is false-, wait, is this true or false?*
>
> Students: *False*
>
> Andrea: *What is it sufficient so that you justify?*
>
> Student C: *One counterexample*
>
> Andrea: *One counterexample, and here [on the whiteboard] we have a counterexample. Is this sufficient in order to assert that this [St143] is false?*
>
> Students: *Yes*

Andrea's teaching made visible one of her assumptions about the sufficient mathematical evidence to disprove a US: providing a counterexample can be considered a justification because the statement says ALL and if there is at least one that does not satisfy, then it is false (dAA1[3], see Chapter 5, Section I.2.1.2).

The teachers' attention to the logical interpretation of statements during their teaching was also observed in other cases. For instance, when explaining why an irrelevant example could not disprove a US (e.g., Andrea's assumption dAA1[9*] and Gessenia's assumption aAGt2; for details see Chapter 5, Sections I.2.2.2 and I.2.2.3); when explaining the status of confirming examples when proving a US (Andrea's assumption dAA9[9]; for details, see Chapter 5, Section I.3.1.1). Notably, Gessenia's teaching exhibited her focus on the logical interpretation of the statements as she drew her students' attention to the status of confirming examples when trying to prove the statement *"Every time we use a bigger number, we will always get more true sentences than for the previous numbers"*[275].

> Gessenia: *How many numbers are here in total (she pointed out the four confirming examples that were written in front of the class)? How many numbers are we talking about? ... Four. But how many are the natural numbers? ... then with these four examples can I already conclude, that this is why the conjecture is true? ... Here Julito says, every time we work with a bigger number, always, he does not make any exception, he does not say not with this one, not with the other one. He says, bigger numbers, always… He says we will ALWAYS obtain more true statements, and we have seen that that is not the case… then what Julito says is not true.*

Gessenia's explanation exhibits the change of her initial assumption that confirming examples are sufficient to prove a US with infinite cases involved. She shows that even though there are four examples that verify the statement, they are not sufficient to conclude that the statement is true. Gessenia is aware that the statement is false and uses the logical interpretation of the statement to draw her students' attention to the insufficiency of a few confirming examples to prove a US.

3. Andrea's emphasis on the number of cases involved in a statement to support the identification of the sufficient evidence to prove a US

Another example is linked to proving universal statements, Andrea's observation about the difference between finite and infinite USs and the importance of considering the size of the set of analysis when proving them. For example, the finite US *"All one-digit multiples of 6 are even"* can be proved by verifying all cases involved in it; whereas the infinite US *"All multiples of 6 are even"* cannot.

During the intervention Andrea learned that an important factor to determine the type of sufficient evidence to prove a US is the number of elements involved in the statement. Andrea also exhibited her new assumptions during her teaching. She drew her students' attention to the "set of analysis" and the number of cases involved in a US when proving it.

> Andrea: *When I am told "all these numbers", what does it refer to? ... Here [task 3: You are given the numbers 0, 3 and 9. Now answer: Is it true that all these numbers are divisible by 1?], which one is my set of analysis? ... Do we also have infinite possibilities here [in contrast to task 1]? ... Here my set of analysis only has three elements… There are only three possibilities that I am going to analyze, that's right... I can verify them all.*

[275] For details on the context, see Chapter 5, Section I.2.2.3, Gessenia's teaching, and/or Appendix A12 for the content of the activity that was the main focus of her teaching.

Through her teaching Andrea made visible the use of her assumption dAA3[1] that as long as the set of analysis is large, examples are not valid to prove a US; if it is a small set, then they are (for details, see Chapter 5, Section I.3.1.1).

4. Gessenia's emphasis on the need to provide evidence when disproving USs

A fourth example of methodological approach that made visible Gessenia's current proof-related assumptions was her focus on highlighting the need to provide a counterexample when disproving a US during her teaching. Recall that at first, during the intervention, Gessenia was hesitant about the sufficiency of a counterexample to disprove a US. However, during her teaching she used two contrasting answers given by her students to emphasize the need to include a counterexample for the statement *"If a distribution is fair, whole and maximal (FWM), then there are zero objects left"* (St45).

> Gessenia: *They [the students who gave an argument without a counterexample] could have used something that proved that what they claimed was true. I need to see, because what you tell me does not explain, it is not evident, I cannot see what you mean exactly. On the other hand, here (Gessenia pointed to Argument A, which included a counterexample), can I see what the group means? ... Yes, because they are showing me a fair, whole and maximal (FWM) distribution and it does not have zero left, therefore this [St45] is not true.*

Gessenia's teaching made visible the use of her new assumption that she needed to see a counterexample; otherwise, it was not evident why the statement was false (aAGt1, for details, see Chapter 5, Section I.2.1.1).

5. Summary of Section III

It is interesting but not unexpected that the most frequently used understanding the teachers promoted during their teaching was their students' understanding of the logical interpretation of statements. It is possible that they regarded it as the most fundamental understanding they could make their students aware of in this context. But in general, these four approaches were likely the most meaningful to them because they were useful, simple and accessible. They were useful as the approaches worked well changing the teachers' initial non-mathematical assumptions. They were simple since they depended on the basic understanding of the logical interpretation of statements. They were accessible because they were simple. Presumably, they decided to include these approaches I used during the intervention as they supported their own understandings.

The teachers' assumptions were visible during their teaching as a consequence of including tasks for kids where the teachers could put them into practice. It is possible that if the teachers had had other opportunities to exhibit further effects of the intervention, then they would have revealed additional assumptions or even some limitations in their current assumptions.

After the intervention and before their teaching the teachers showed their skepticism of whether it was possible that their students could successfully engage in proof-related activities. At the beginning of their lessons the teachers struggled to engage their students in proof-related discussions; however, my company during their teaching and a few of my interventions in their classrooms seem to have helped them gain confidence that it was indeed possible.

The teachers' use of these methodological approaches shows the impact the intervention had not only in their proof-related assumptions, but also in their teaching.

IV. Summary of Chapter 6

The teachers' ability to engage in dis/proving improved during the intervention as was shown by their development of assumptions more in line with mathematical practice as the intervention went on. In comparison with proving USs, the teachers showed more confidence when engaged in disproving USs. The insights they gained about disproving USs and specifically about the characterization of counterexamples contributed to, for example, Andrea's extension of her understanding to accept non-minimal valid justifications to disprove USs (see Section I.7). In contrast, it was not always simple for them to successfully engage in proving infinite USs, although all the teachers became aware of the need to provide proofs that encompassed all the cases involved in the statement during the intervention. This means that when they did not produce proofs it was not necessarily because they had an *empirical justification scheme* (Harel & Sowder, 1998, 2007). It is clear that the teachers' perceptions of proof were more aligned with the mathematical perspective by the end of the intervention. I agree with G. J. Stylianides and A. J. Stylianides (2020) that this is an important factor that should not be overlooked when investigating individuals' justification schemes, especially as I was interested in evaluating the development of the teachers' proof-related assumptions. Nevertheless, whether a teacher and others can or cannot produce proofs is a matter that involves additional factors. G. J. Stylianides and A. J. Stylianides (2020) have pointed out some factors to take into consideration. In reference to my intervention, I conclude that an important factor might be the three teachers' lack of knowledge of methods to prove infinite USs (e.g., use of variables to represent any element in the set of analysis, mathematical induction), which are tools to tackle generality. In addition, there might be cases where teachers (like Lizbeth in my research) and others are aware that one counterexample is sufficient to disprove a universal statement. However, as they struggle and do not succeed in finding one counterexample, they might conclude that a false universal statement is true (e.g., Zaslavsky & Ron, 1998), unlike Lizbeth who was aware that to be true, a universal statement should not have counterexamples (see Section I.7). The teachers' understanding of the nature of proving also affected their ability to prove and disprove existential statements. The teachers proved existential statements by showing at least one confirming example, which they recognized as sufficient evidence. This understanding led Andrea to come up with an original approach to disprove existential statements. She concluded that a "some-statement" is false if it is impossible to find confirming examples for it (see Section I.7).

My work with the teachers has not only shown that the teachers improved their proof-related assumptions. More importantly, it has shown that the assumptions they hold are often linked to each other and this has an important influence on their development. In some cases, once an assumption was changed, it had an impact on other assumptions. For example, Andrea and Gessenia initially accepted a number of confirming examples to prove a universal statement. Later, they assumed that confirming examples (in the form of specific calculations) did not qualify as sufficient mathematical evidence to prove (in general). It had an influence on their disproving of USs as they used that assumption to discard counterexamples to disprove a US (see Section I.6). This shows the close connection the teachers may establish within their set of assumptions and because of the interwoven relation of their assumptions, attention should be given to them in a conjoint way. Further, the teachers' assumptions show that even though they might use mathematical assumptions, the reasons that support their assumptions might not be mathematical (see Section I.4). This suggests that attention should be given not only to the assumptions the teachers and others develop, but above all to their warrants and what nature they have.

Several features of the intervention supported the development of the teachers' mathematically aligned proof-related assumptions. One of them turned out to be fundamental, that is the teachers' attention to the logical interpretation of single-quantified statements. This feature was fundamental as it underpinned several assumptions. For their assumptions the teachers relied on prior assumptions developed during the intervention, which ultimately relied on their understanding of the logical interpretation of the statements. Its importance was also visible during the teachers' teaching as they applied it to promote their students' proof-related activity. It does not discard that other features were also important. In fact, features complemented each other. For example, challenging the teachers' assumptions turned the teachers' attention to relevant aspects that supported the refinement of their current assumptions. Provoking cognitive conflicts was one important way to challenge the teachers' assumptions. Certain types of statements triggered cognitive conflicts for one teacher, although not for the others. Using different types of statements supported the realization of cognitive conflicts and in some other cases the identification of fundamental initial assumptions the teachers used. The gradual development of key proof-related concepts such as "proof", "some" and "counterexample" allowed the teachers to refine their own assumptions about the concepts at their own pace and make sense of them. Provoking cognitive conflicts also supported this. The expectation for mathematical justifications was an overarching feature that supported the development of the discussions, the teachers' attention and reflection on mathematical aspects and the establishment of their assumptions. The feature related to the size of the group of teachers who participated in the intervention was also important. It allowed the teachers to feel free to openly share ideas, show disagreements and learn from each other.

There were also some aspects of the intervention that might explain the failure to change some of the teachers' assumptions. These are particularly related to the case of negations and "no-statements". For example, introducing negation with rules was an obstacle for the teachers' search for the reasons that explain why the rules worked as they did. A lack of access to some of the teachers' fundamental assumptions impeded addressing and overcoming root problems when they first emerged. A lack of awareness of the direct relation between language and proof-related issues made it difficult to find out where the teachers' struggles stemmed from and properly react to them when needed. There was limited usage of single-quantified statements in their varied forms, and a lack of discussion and explicit establishment of connections among them. This meant that the teachers did not have many opportunities to use their mathematically-aligned assumptions for the development of assumptions about equivalent statements.

During their teaching some of the teachers' assumptions were visible. In particular, it was interesting to see that they applied some approaches I used during the intervention to their lessons. For example, they drew their students' attention to the logical interpretation of the statements in discussion, which implied their emphasis on the quantifiers and the number of cases involved in the set of analysis of the statement. The teachers suggested irrelevant and confirming examples to support their students' rejection of cases that did not qualify as counterexamples for false USs. They also drew attention to the need to provide mathematical evidence to dis/prove a statement and discussed what that evidence consisted in. The usage of these approaches did not only make visible their own current assumptions, but also the relevance they ascribed to them. Both the features of the 2018-intervention that supported changes and those that hindered them that I consider in this chapter foreshadow principles that I include in Chapter 7.

My work points out the need for research that has a focus on developing a global view of teachers and others' proof-related assumptions. This involves for example investigating the development of their assumptions about proving and disproving, universal and existential statements, whether those assumptions are related, what their groundings are and the interpretation of key proof-related terms, like quantifiers, simultaneously. I do not claim that it is not important to have details on specific issues about their proof-related activity; however, instead of cutting up these analyses, they should integrate different relevant aspects that allow a more complete view of teachers and others' proof-related assumptions.

Chapter 7: Implications for Future Teacher Development Interventions

"The only complete safeguard against reasoning ill, is the habit of reasoning well; familiarity with the principles of correct reasoning; and practice in applying those principles."

(John Stuart Mill, as cited in Epp, 2020, p. 146)

The quotation above points to "principles of correct reasoning", but what are those principles? Some hints in that direction have been given in Epp (2003, 2009) and Durand-Guerrier et al. (2012). Inspired by their recommendations, I adopted, adapted and explored some of them in my research. Here I offer some suggestions that arose from my reflection on the outcomes (both strengths and limitations) of my work, and which may support engagement in correct reasoning.

Some of the recommendations I make here reflect successes during my 2018-intervention and others are based on my failed attempts during the intervention. They are intended for future interventions with teachers with a focus on their capability to engage in PfBT. These recommendations can be seen as the final design principles arising from my design-based research, and answer my third Research Question:

- RQ3: What design principles for a teacher development intervention focused on Proof-Based Teaching (PfBT) in primary schools can be abstracted from two cycles of such interventions?

The seven design principles in Part I are general, and the two in Part II are specific to blocks about no-statements and negations.

I. General design principles

1. Principles about the content of the intervention

I recommend dividing the intervention into two main parts: the first part should have a focus on a specific mathematical content, and the second part on the development of mathematical reasoning principles to be formulated and explained by the teachers. The mathematical content in the first part should be simple. This allows the teachers' attention to be drawn mainly to the mathematical reasoning principles to be developed during the second part of the intervention. The mathematical statements to be used during the second part of the intervention should be framed in the mathematical content developed during the first part of the intervention. Again, the main goal is not to distract the teachers' focus from the reasoning principles with difficulties related to the mathematical content.

1.1. First part of the intervention: The mathematical content

An important strategy to support the teachers' development of mathematical reasoning principles is provoking cognitive conflicts (see Section 7.1 below). Chinn and Brewer (1993) recommend helping students construct needed background knowledge with that purpose. In this context, the mathematical content section of the intervention is expected to help the teachers construct knowledge in order to support the cognitive conflicts with their non-mathematical assumptions.

Given that the main goal of the intervention is to foster the development of mathematical reasoning principles, the mathematical content for the first part of the intervention should be easy enough that is does not become an obstacle for this main aim. For example, in the

2018-intervention I focused on division and divisibility in a PfBT context because: (1) I was working with primary school teachers, whose mathematical background is usually not very strong; however, division and divisibility is a mathematical content that can be straightforwardly understood in a PfBT context, not only by teachers, but also by young children (see Vallejo-Vargas & Ordoñez-Montañez, 2014; Vallejo-Vargas & Ordoñez-Montañez, 2015). (2) The advantage of working in a PfBT context is that the knowledge construction relies on a small number of initial definitions/axioms/principles, which facilitates direct proving, the first approach the teachers use to prove universal statements that involve infinite cases. (3) It is important that all participants share common groundings in the mathematical content; otherwise, it might be confusing for the teachers to have different sources for the rationales they use to explain why a statement is true or false during the second part of the intervention. For instance, in the case of the 2018-intervention, that common basis was the notion of fair, whole and maximal distributions.

When selecting a mathematical content, a crucial point to consider is choosing content where it can be easily established what the toolbox (Netz, 1999) of established knowledge is. The teachers should know what the foundations or initial assumptions are related to the content. There is a need to be explicit about the key notions, axioms, definitions, principles, etc. that belong to the toolbox to be shared by all members of the classroom community. These elements are linked to the first component of the definition of proof given by A. J. Stylianides (2007) in a school context. They are the "statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification" (p. 291).

Using a new context (a proof-based context) to learn "well-known" content allows the teachers to re-learn the content by using only the elements available in the toolbox at that moment. This restricts their reasoning and proofs to those elements so that they become consciously aware of the theorems they make as well as of the principles that guarantee those theorems. Additionally, having the teachers teach the same content to their students in the same proof-based way facilitates the teachers' design and teaching development of their lessons. They would not have to worry about how to organize their teaching on their own as they can count on the design used during the intervention. It can be taken as a reference for the teachers to come up with appropriate re-arrangements of the tasks.

1.2. Second part of the intervention: The mathematical reasoning principles

The second part of the intervention has a focus on the development of mathematical reasoning principles. Among those, I recommend those related to the discussions I included in Table 31 (see Section 2 below), which I grouped by blocks (first column) that include a certain number of discussions (second column). I specify the understandings that are expected to be promoted in each discussion (third column). Lastly, I include the type of statements in the order they should be presented in order to promote each of those understandings (fourth column).

I suggest organizing this part of the intervention into four main blocks. The first has a focus on universal statements (USs), the second on existential statements (ESs), the third on "no-statements" (NO-Ss) and finally the fourth on negations.

A rough time duration for an intervention that includes all of these blocks is about 24 hours, which is the approximate time I had for the 2018-intervention. It is important to consider that the first block, which is the basis for the others, may last longer than other blocks. Other blocks might take longer than the time I used for them during the 2018-

intervention because of the additional suggestions I make here. For example, Block II has a focus on existential statements; however, the 2018-intervention did not include a discussion to establish a linkage between existential statements in different forms (e.g., "some-statements", "there-exist-statements", "there-is-statements"). That is suggested as a new part of Discussion 6, focused on the understanding of the logical interpretation of existential statements.

The time needed for each discussion will also depend on the teachers who participate in the intervention, and how willing they are to express their thinking, so the timing of the intervention cannot be specified exactly in advance.

2. Principles about the order of the blocks in the second part of the intervention

"A reasonable argument can be made that the most important form of statement in mathematics is the universal conditional statement."

(Epp, 2020, p. 113)

"... most students find it easier to understand and construct disproofs by counterexample than to understand and construct even simple direct proofs."

(Epp, 2009b, p. 314)

Most of the statements used in mathematics are universal conditional statements and the quantification is sometimes left implicit (Epp, 2020), which might hinder their understanding. Furthermore, as the second quote states and my personal experience endorses, disproving universal statements is easier to understand than proving them (Epp, 2009b). These are the main reasons why I opted to begin the 2018-intervention with universal statements (Section 2.1) and suggest the same for future interventions. I suggest then continuing with ESs (Section 2.2), proceeding to NO-Ss (Section 2.3) and finishing with Negations (Section 2.4). In each of the following sections I explain why I recommend this structure for the design of future interventions.

2.1. Begin with Universal Statements

I suggest beginning a future intervention with USs (Block I in Table 31). I concur with Epp (2009b) that disproving USs is easier to understand and perform than proving USs. This is the main reason why I suggest (and opted) to begin the proof-related discussions with disproving universal statements and explicit discussions about related aspects. A discussion about disproving USs should come before fostering understanding of proving USs. Nonetheless, my suggestion is that before directly engaging in disproving USs (Discussion 2), the focus of discussions should be on understanding the logical interpretation of USs (Discussion 1, for details see Section 3.1 below). In order to create a need for proof, I suggest that Discussion 3 focus on understanding the status of a statement, which involves for instance becoming aware of the difference between a conjecture and a mathematical truth. Discussion 4 is intended to promote understanding of what is involved in establishing truth of and proving a US. Finally, Discussion 5 is a discussion about the difference between a US and its converse. I place this discussion at the end of block I because it is important that the teachers are aware of when a US is true or false, before engaging into noticing a distinction between a US and its converse.

Table 31

Block	Discussion #	Understanding to be promoted	Statement
Block I: Understanding Universal Statements	Discussion 1	Understanding the logical interpretation of Universal Statement (USs)	UASs
			UNSs
			UCSs
	Discussion 2	Understanding what is involved in establishing falsity and disproving a US	UASs
			UNSs
	Discussion 3	Understanding the status of a statement	
	Discussion 4	Understanding what is involved in establishing truth and proving a US	UASs
			UNSs
	Discussion 5	Understanding the difference between a US and its converse	UASs
Block II: Understanding Existential Statements	Discussion 6	Understanding the logical interpretation of an Existential Statement (ES); Understanding what is involved in establishing truth of and proving ESs	EASs and ENSs
	Discussion 7	Understanding disproving ESs	ENSs
			EASs
Block III: Understanding "No-statements"	Discussion 8	Understanding the logical interpretation of a No-Statements (NO-S)	
Block IV: Understanding Negations	Discussion 9	Understanding the connections between USs and ESs through Negations	

Recall that, within the class of USs, I consider three types of statements; namely: universal affirmative statements (UASs), universal negative statements (UNSs) and universal conditional statements (UCSs) (see Chapter 3, Section II for details). I suggest that they be addressed in the order laid out in the next three sections.

2.1.1. Begin with Universal Affirmative Statements

From my experience with the 2018-intervention, the teachers' initial interpretation of affirmative "all-statements" (i.e., a statement of the form *"All X are Y"*) was already close to a mathematical interpretation and details were refined during the intervention.

My suggestion is that an intervention with a proof-related focus begins with an affirmative "all-statement". The goal would be to make the teachers aware of the logical interpretation of "all-statements" (see Section 3.1 below), and later link this insight to the understanding of UCSs. That is, the understanding of UCSs is aimed to be built on the understanding of "all-statements" or equivalent statements (e.g., "every-statements" and "each-statements").

Besides that, using first a mathematical statement situates the teachers in the mathematical framework. They should be expressly told that the statements are to be evaluated mathematically; however, the statements should not only include statements about mathematics content (see Section 4.1 below).

2.1.2. Continue with Universal Negative Statements

Affirmative statements are more common than negative ones. For a future intervention I suggest introducing UNSs during Discussion 1, as a contrast with UASs. Specifically, I recommend using an affirmative and a negative "each-statement" that share the respective sets involved and that are framed in an imaginary context (e.g., *"Each «bogui» is a «fantaslopitocus»"* and *"Each «bogui» is not a «fantaslopitocus»"*). For the comparison I recommend using a UAS that has been analyzed in a previous discussion. Considering both affirmative and negative statements in the same discussion makes the difference more explicit and evident.

I recommend using a negative "each-statement" instead of a negative "all-statement" because during the 2018-intervention Andrea confounded *"All X are not Y"* with *"Not all X are Y"*[276]. In addition, she initially interpreted the latter as *"Some X are Y"* as she used a *semantic substitution* of "not all" with "some"[277]. However, Andrea did not seem to struggle with a UNS of the form *"For every natural number n, F(n) is not P"*[278]. This suggests that the form of the statement had an influence in its interpretation (see Section 3.1.4 below). For future interventions I suggest avoiding the possibility that the teachers might resort to *semantic substitutions* at this point, given that it might divert the focus of the discussion towards a discussion about existential statements. When using a statement of the form *"Every X is not Y"* in English, a participant might assume that it is the same as *"Not every X is Y"*, as Andrea did with "all-statements". However, in Spanish there is no simple way to place the negator "not" in front of an "every-statement" as can be done in English. Hence, it is possible that the case of a negative "each-statement" could work better in English and that is why I suggest engaging in this discussion with both an affirmative and negative "each-statement". My expectation is that the teachers mobilize their insights for the affirmative "each-statement" to make sense of the negative "each-statement". The purpose of using imaginary statements is to focus on general reasons (see Section 4.1 below). Later the understanding of negative "each-" or "every-" statements might support the understanding of negative "all-statements".

2.1.3. Finish with Universal Conditional Statements

Once UASs and UNSs are properly understood, I recommend continuing with UCSs. I suggest including first an imaginary affirmative "if-then-statement" and making connections with the universal statements that were seen previously. For example, a UAS from an earlier discussion can be changed into its "if-then" form and used to introduce conditional statements (e.g., *"If it is a «bogui», then it is a «fantaslopitocus»"*). This should allow exploration of the teachers' initial assumptions related to the interpretation of CSs, independently of the truth value of the statement, but also reveal if they link CSs with USs.

For a subsequent discussion I suggest considering two mathematical "if-then-statements", that both should share the same set of analysis and conclusion set, but one of them should be affirmative and the other one negative. The rationale for this suggestion is similar to that used for the introduction of UNSs (see Section 2.1.2 above).

It is important to ask the teachers whether the examples of "if-then-statements" are equivalent to universal statements, as a way to notice a link between statements. That might allow them to transfer their understanding of USs to UCSs. In that respect, I concur with Epp (2020) that "[t]he crucial point is that the ability to translate among various ways of expressing universal conditional statements is enormously useful for doing mathematics and many parts of computer science" (p. 3). Be aware that teachers may overgeneralize that all conditional statements are equivalent to universal statements. In Section 7.1.4 I include a suggestion of how to avoid that assumption.

[276] For details, see Chapter 5, Section III.2.2.
[277] For details, see Chapter 5, Section I.4.
[278] For details, see Chapter 5, Section III.2.1.

2.2. Continue with Existential Statements

I recommend continuing Block II with a focus on ESs. Begin fostering the understanding of the logical interpretation of ESs (Discussion 6), where related assumptions are expected to be changed to fit a mathematical perspective. This understanding will facilitate proving ESs as it directly relies on a proper understanding of the logical interpretation of an ES. The teachers should understand at this stage that "some" and "all" are not mutually exclusive. Disproving ESs can follow and be understood through the understanding of proving of ESs (Discussion 7). Tabach et al. (2010a) claim that "[e]xistential, false statements are rarely addressed in school mathematics" (p. 1086). In Section 7.4.2 (below) I suggest an approach that has that goal.

2.3. Proceed with No-Statements

My suggestion is to continue with Block III and the understanding of NO-Ss. The approach I suggest for promoting understanding of NO-Ss relies heavily on understanding USs first, so I included NO-Ss later than USs. I cannot see any reason to discuss NO-Ss before ESs; however, it might possibly bring positive outcomes. It is worth investigating whether the order of Blocks II and III affects the development of the teachers' understanding.

2.4. Finish with Negations

I suggest finishing the discussions with the negation of single-quantified statements. This discussion should allow the teachers to see more clearly how USs and ESs are linked, and so it must come after discussions of USs and ESs.

3. Principles about the content of each block

All the blocks in a future intervention should aim at a common core understanding. Each block should include first a discussion on the logical interpretation of statements (Section 3.1). This recommendation applies to each type of statement that is the focus of the three first blocks (see Table 31). In addition, the blocks should emphasize the status of statements (Section 3.2).

3.1. Before engaging in proof-related discussions, make sure the participants understand the logical interpretation of the statement

"I told you before that here [#4] it can be both true and false? I was wrong. It's only true. […] I didn't understand the statement properly. I was sure that they are talking about all numbers, as I wrote here [#3]," Debby said.

(Buchbinder & Zaslavsky, 2009, p. 231)

This quote from Buchbinder and Zaslavsky (2009) underscores the need to understand statements before engaging in other activities such as dis/proving. Debby, the 10th grade student in Buchbinder and Zaslavsky's study, explicitly pointed out this aspect. At the beginning, she did not properly understand statement #4, an existential statement[279], as

[279] In Buchbinder and Zaslavsky's (2009) research, the authors used a list of six different quantified statements, so that the students could determine whether they were true or false.

she thought it referred to all numbers, like the previous statement #3, a universal statement, and this led her initially to ascribe a different truth value to statement #4.

In Chapter 3 (Section II.1) I explained what I mean by understanding the logical interpretation of SQ-statements. In short, it involves understanding the close relation between the three main elements in a SQ-statement: the “set of analysis” and “the conclusion set”; the elements that belong to the set of analysis and that the quantifier applies to (all or some, according to the quantification); and the claim made about those elements, according to the conclusion.

The focus of the 2018-intervention was SQ-statements as my main target was in-service primary school teachers. It makes sense to discuss SQ-statements before engaging in multiply quantified statements. In this respect, Dubinsky, Elterman and Gong (1988) recommend that “[b]efore passing on to higher level quantifications it is important to note that the encapsulation of single-level quantifications is critical for working with several quantifications” (p. 48). My attention is focused on developing teachers’ understanding of the logical interpretation of SQ-statements, although I am aware that interpreting mathematical statements might be a non-trivial work even for undergraduate students (e.g., Dawkins, 2017; Ferrari, 2004; Durand-Guerrier, 2003; Epp, 2003).

Whether a statement is given in its formal or informal version, understanding its logical interpretation should be, in my view, a first step prior to discussing further issues about it. I concur with Selden and Selden’s (1995) suggestion that unpacking the logical structure of informal statements is important when validating a proof since, as they claim, “who cannot reliably unpack the logical structure of informally stated theorems, also cannot reliably validate their proofs” (p. 130). I connect unpacking the logical structure of informal statements with understanding the logical interpretation of a statement. I contend that being able to unpack the logical structure of informal statements entails understanding the different logical structures statements can have. This necessarily involves understanding their logical interpretation.

Given that it is crucial to understand the logical interpretation of statements before engaging in other related activities about them, I suggest that each of the three first blocks (see Table 31 above) begins with it. This is a fundamental understanding and serves as the basis for the development of proof-related insights. Many issues in relation to this type of understanding stem from the everyday meanings that individuals use for quantifiers. They bring their personal usages of those quantifiers to the discussions, which might hinder the development of their understanding in a mathematical context, as I observed in my work (see Chapter 5, Sections I.1, II.1, III.1.1, and III.2.2.1). This is mainly related to their non-mathematical identification of what elements from the set of analysis the statement refers to. Thus, the teachers failed to engage in processes that depended on those interpretations; for instance, when proving or disproving SQ-statements. This kind of understanding is what I call here “understanding the logical interpretation of a statement” (I target these understandings in Discussions 1, 6 and 8 in Table 31).

Understanding the logical interpretation of a statement was essential for the teachers who participated in the 2018-intervention. Clear evidence of this is their own initiative to draw their students’ attention to this understanding before identifying what is involved in proving or disproving statements (see e.g., Chapter 6, Section III.2).

Some concrete recommendations for developing this type of understanding are: use examples of statements framed in different contexts, use statements with different truth values, promote the use of diagrams as alternative ways to represent the logical interpretation of a statement, use statements in different forms, verify that examples of elements from the set of analysis can be given, and request comparisons among statements. I will address each of these in the following sub-sections.

3.1.1. Shift the content of statements

I noticed in the 2018-intervention that the scope of an assumption could be tested by changing the content in which the statement was framed and paying attention to whether the assumption changes or the teacher hesitates about it. For example, Andrea used her initial non-mathematical assumption that *a true UAS implied that its converse was false* with statements with different content during the 2018-intervention (see Chapter 5, Section I.1.1.1). This means that her assumption was independent of the content in which the statements were embedded and as such it was resistant to change.

The contents I considered during the 2018-intervention were: daily-life or familiar, mathematical and imaginary. In a new intervention I also suggest including the abstract context, in which terms the reasoning principles can be ultimately formulated (see Chapter 4, Section II.2.1, Stage 1.3, for details on types of statements according to their content).

3.1.2. Vary the truth value of statements

True statements are not strictly required as starting points in order to determine the logical interpretation of a statement or to find equivalent statements[280], unlike when making sound valid mathematical inferences (Baggini & Fosl, 2010).

My experience with the 2018-intervention taught me that teachers tend to confuse, for example, the truth relationship of the two sets involved in a statement with the claim made in it. An illustration of this confusion is Lizbeth's attempt to represent the logical interpretation of the familiar statement *"All dogs are dinosaurs"* during the 2018-intervention. She used a representation of two disjoint sets, the set of all dogs and the set of all dinosaurs as she claimed that dogs were not dinosaurs and so the sets should be separated. This implied that her attention was placed on her everyday knowledge, which coincides with the truth relationship between the two sets involved in the statement. This kind of assumption mostly arose when the given statement stated something against the teachers' everyday knowledge; that is, the statement stated something that differed from the truth relation between the sets involved in the statement.

Including a harmonious familiar statement (see below Section 4.1 for details) can support the understanding of the logical interpretation of a statement and avoid conflicts as that Lizbeth faced during the 2018-intervention. However, it is also desirable that teachers have opportunities to reflect on the logical interpretation of false statements so that their understanding can be tested.

3.1.3. Use diagrams to promote understanding

During the 2018-intervention I introduced Euler diagrams as an optional way to represent the logical interpretation of a universal statement. Using diagrams allowed me to have access to some of the teachers' initial assumptions (see Section 4.2.2 below); however, my lack of emphasis in their usage might possibly have prevented the teachers from using this powerful tool to reason about mathematical statements. In particular, diagrams can be used to test the validity of arguments (see e.g., Epp, 2020, pp. 151–153); however, that involves understanding the diagrammatic representation of the logical interpretation of a statement. Moreover, diagrams can support subsequent discussions about inferences. The

[280] Here I use "equivalent statements" as statements that state or convey exactly the same thing.

use of diagrams can facilitate visual comparisons of the logical interpretation of statements. I suggest including tasks where teachers are requested to represent the logical interpretation of different kinds of statements, not only universal statements, as I did in the 2018-intervention. Representations can be used as a means to promote further understanding of statements. To what extent visual representations can be used to support reasoning about statements is an interesting topic that I would like to investigate further in the future.

3.1.4. Use different forms of statements

During the 2018-intervention I observed that the form in which SQ-statements are presented influences their interpretation and therefore proof-related activities with them. My recommendation is to draw the teachers' attention to the different forms in which SQ-statements can be presented.

In terms of the quantifier involved in a SQ-statement, most of the universal statements I used during the 2018-intervention were either "all-statements" or universal "if-then-statements". I supposed that the teachers interpreted other universal quantifiers such as "every" or "each" in the same way as they understood "all" and that is why I did not intentionally include USs with those quantifiers; however, the 2018-intervention showed me that universal negative statements of the form *"All X are not Y"* and *"For every x in A, P(x) is not Q"* were interpreted differently by the teachers (see Chapter 5, Section III.2). In the same vein, the teachers found differences among equivalent existential quantifiers such as "some" and "there exist" or "there is" (see Chapter 5, Section II.1). For example, for the teachers, disproving "some-statements" was not as simple as disproving "there-is/are-statements" and "there-exist-statements" (see Chapter 5, Section II.2).

Different equivalent quantifiers in mathematics might have different meanings in an everyday context. Attention should be paid to this issue. For example, Freeman and Stedmon (1986) examined the case of the universal quantifiers "all", "each" and "every". They explain that these words have both a determinative and a quantifying function, and "[u]niversal quantification demands the coordination of quantifying and determining" (p. 30). Furthermore, unlike "each" that modifies a singular noun phrase, "all" modifies a plural noun phrase. For example, we say "Each person is …" and "All people are…". This is, as Vendler put it, "no mere caprice of grammar: it is indicative of a difference in the very meaning of these words" (as cited in Freeman & Stedmon, 1986, p. 35). In the context of mathematical/logical quantification all these forms are equivalent, but in natural language they are not: "all" is collective, "each" is distributive, and "every", in Freeman and Stedmon's terms, "may be regarded as a distributive which has collective ambitions" (p. 36). Another relevant example is the indefinite article "a", which is used in the *mathematics register*[281] (Halliday, 1978) as an implicit universal quantifier. Durand Guerrier et al. (2016) explain that "[a] well-known difficulty is related to the meaning of the article 'a' that can either refer to an individual, a generic element, or an implicit universal quantifier" (p. 89).

In addition, during the 2018-intervention the teachers engaged in a discussion about the UAS *"Zero is divisible by every natural number different from zero",* which has a form different from the other statements we had seen. This statement turned out to be especially puzzling for Andrea, who assumed that the only element in the set of analysis was zero. There are two reasons why I believe it is important to consider the same statement in a

[281] For details on "mathematics register", see Chapter 3, Section I.2.3.

future intervention. First, the quantification is not evident in the way the statement is formulated. Its form is quite different from that of other universal statements. Second, the statement itself is relevant, as it is a theorem the teachers came up with during the first part of the 2018-intervention.

Considering the previous points, I suggest including statements in several different forms; for example, universal statements can be presented as "all-", "every-", "each-" statements and "if-then-statements" can also be presented in alternative ways (e.g., *"If then Y"*, *"Y, if X"*, *"when X, then Y"*). It is possible that teachers might make different assumptions about them. Connections among them should be made by basing understanding of harder-to-grasp statements on understanding of easier ones. Identifying the form of expression for which their related assumptions are closer to mathematical convention can provide a good opportunity to establish connections with other equivalent statements with different forms and transfer related assumptions once the equivalence is clear. That is, certain forms of statements might be used as a bridge for other forms of equivalent statements and mathematical assumptions. An appropriate understanding of the logical interpretation of SQ-statements can support, for instance, the disproving of a "some-statement" by relying on its equivalent "there-exist-statement". The latter can be used as a bridge statement for the disproving of a "some-statement".

3.1.5. Make sure that the participants can exemplify elements from the set of analysis

My recommendation is to include tasks that target the identification of elements in the set of analysis. During the 2018-intervention Gessenia struggled with this type of task when the elements were different from simple natural numbers, as in most of the cases we analyzed. For example, if the set of analysis was a subset of the pairs in $\mathrm{N} \times \mathrm{N}$, as in the statement *"All divisions are exact divisions"*, then the teachers sometimes struggled to identify the elements of the set of analysis. This was particularly the case with the statement *"If A is divisible by B, then B is divisible by A"*. Explicitly asking for elements in the set of analysis that are pairs, before requesting the teachers to engage in further discussions about the statement, might avoid distractions from the main goal, that is understanding its logical interpretation.

3.1.6. Request comparisons between statements

To better understand the logical interpretation of a statement, it is useful to include, for example, both affirmative and negative statements with the same set of analysis and conclusion set, in order to compare their logical interpretation. Representations can be used with the purpose of comparing the logical interpretation of the statements. A request to explicitly establish commonalities and differences among them can support this goal.

3.2. Emphasize the status of statements

Individuals tend to overgeneralize patterns and presume they are generally true because of their recurrent behavior. In particular, this is the way many scientific fields develop their theories (Reid & Knipping, 2010). Nonetheless, mathematics is different in this sense and supporting the teachers' awareness about this difference is crucial. Discussion 3 has a particular focus on this topic (see Table 31).

A clear distinction should be made about the status of statements. It is important to clearly establish the status of a statement (i.e., whether it is a conjecture, a true statement, a false statement) once it has been proved, disproved or even when a proof for it has not yet been found.

Using the expression "mathematical truth" to refer to a statement that had been proved to be true turned out to be a good way to establish a distinction between conjectures and theorems during the 2018-intervention. Moreover, the teachers learned that a valid justification was in need for a statement to be "upgraded" from a conjecture to a mathematical truth. In Table 31, I suggest focusing on *Proving USs* first in Discussion 4. Hence, to illustrate a *mathematical truth* in Discussion 3 I recommend using a true conjecture that the teachers proved[282] during the first part of the intervention.

The statement *"For every natural number $n > 0$, the expression: $1+1141n^2$ does not produce a perfect square"* with a hard-to-find counterexample directs attention towards an important and determinant factor when discussing the status of US-conjecture statements; namely, justifications. It shows that nothing can replace a mathematically valid justification in a discussion about whether a conjecture is true.

During the 2018-intervention I was not explicit about the status of false-conjecture statements and this had an impact on Gessenia. She explained to her students that after a conjecture was disproved, it remained a conjecture. This showed me that an explicit discussion about the status of a conjecture that is proved to be false – a *refuted conjecture* – should be considered in future interventions.

In the same line, I also suggest including examples of *open conjectures* with the aim of pointing to the status of conjectures for which no proof or disproof has been found yet. For instance, during the 2018-intervention I used the strong and the weak Goldbach's conjectures to show that in spite of centuries of research a conjecture could still be open (e.g., Goldbach's strong conjecture), while others are only proved after many years (e.g., Goldbach's weak conjecture). Furthermore, adding that the Goldbach's weak conjecture was proved by a Peruvian mathematician was an interesting fact in our context. During the 2018-intervention I also talked about the (now) six millennium problems (as examples of open conjectures) and the prize offered to those who could solve them. This was particularly thought-provoking for the teachers. In such discussions there is an expectation that the teachers can see that proving a statement is a non-trivial effort and that finding a mathematically valid justification is crucial.

4. Principles about the content of the discussions

Each discussion should encompass different sub-discussions, all of which serve the main goal of promoting the underlying understanding. Each sub-discussion could be presented as a slide. Each has a certain focus, while in some of them there is an expectation that the teachers formulate a principle. There is nothing special about the number of sub-discussions each discussion should have. It depends on the understandings that are to be developed within such discussion. Here I include three general principles: using different types of statements to facilitate the establishment of general mathematical reasoning principles (Section 4.1), beginning each discussion by identifying initial assumptions (Section 4.2), finishing each discussion explicitly establishing the new assumptions (Section 4.3).

[282] My expectation is that most of the conjectures the teachers come up with during the *first* part of the intervention are proved by direct proving, that is, basically relying on the given definitions.

4.1. Use different types of statements to facilitate the establishment of general mathematical reasoning principles

Some researchers have given some hints about the role that the organization of tasks plays in the individuals' performance. The content and organization of the sequence of problems affect the display of logical skills, in particular when solving syllogisms (see e.g., Hawkins et al., 1984). Epp (2003) recommends introducing mathematical reasoning principles with the support of statements whose daily-life interpretation matches that from a mathematical perspective. She claims that this can "motivate students to accept the reasonableness of the principles of mathematical reasoning" (p. 895).

I had some of these suggestions in mind when I designed the 2018-intervention; however, the insights I gained from the intervention along with further reflection on the literature leads me now to recommend a somewhat different task organization to support the formulation of mathematical reasoning principles (see Table 31).

Before I suggest an organization for the statements in a discussion, I introduce some terminology. In addition to the "familiar", "imaginary" and "abstract" statements I described earlier (see Chapter 4, Section II.2.1, Stage 1.3); here I describe two new types of statements that are needed to make sense of the organization: "harmonious statements" and "problematic statements".

Harmonious statements

A *harmonious statement* is a mathematical or familiar statement whose common interpretation agrees, matches or is "in harmony" with its mathematical interpretation and the main aim of the task or discussion. For example, if the focus of a discussion is on the logical interpretation of a US, a harmonious statement should be a *true* US since there is no conflict between the truth of the statement and what it is claimed in it. Notably, an example of a harmonious statement for a discussion on the logical interpretation of affirmative "all-statements" would be the familiar statement *"All humans beings are mortals"*. An example of a harmonious statement for a discussion on affirmative conditional statements (CSs) could be the mathematical CS *"If a natural number is smaller than 2, then it is smaller than 5"*. If the focus is on disproving USs, a harmonious statement would be a false US. The familiar UAS *"All lawyers are teachers"* is an example of a harmonious statement in such a context. A harmonious statement for a discussion on the relation between a UAS and its converse, for which it is usually assumed that a true UAS implies that its converse is true, could be *"If a number is even, then it halves"*.

A harmonious statement is "harmonious" in theory. This means that a participant might not realize that a statement is harmonious in practice (e.g., in the case that the participant cannot determine whether the statement *"If a natural number is smaller than 2, then it is smaller than 5"* is true); however, planning requires awareness of the statement's "harmonious" nature. The design of the intervention is based on this awareness despite the possibility that it might be unnoticed by some participants.

Whether a statement qualifies as harmonious or not does not necessarily require the participant's familiarity with the content in which the statement is framed. For example, the mathematical statement *"If a number is palindrome, then it is a square number"* is a harmonious statement in the context of discussing false UCSs even though the participant might not know what palindrome or square numbers are. The participant would probably fail to determine that the statement is false; however, it is a harmonious statement in the sense that it is indeed a false statement and so discussing a disproof for it is possible. It is desirable, however, that the content of the statement is compatible with the participant's background knowledge so that it is not an obstacle to accessing her/his current assumptions.

Problematic statements

A *problematic statement* is a familiar or mathematical statement that is introduced to create uncertainty in relation to an assumption that is held by a participant or that was just established. Problematic statements aim at challenging the participants' current assumptions. For example, the problematic (familiar) statement *"If it is a person, then it is a mortal"* supported Gessenia's change of her initial assumption, that a UAS implies its converse, through her realization that a true UAS may not imply its converse[283].

Discussing problematic statements involves an expectation for a reconsideration and reformulation of an assumption. Furthermore, a statement that is problematic for one participant may not be problematic for another. For example, the true ES *"Some numbers divisible by 4 are even numbers"* is a problematic statement in the context of a discussion for the logical interpretation and proving of ESs. It is a problematic statement in such a context because of the typical everyday interpretation for "some" (i.e., "some, but not all"). This example of problematic statement was particularly relevant for Andrea during the 2018-intervention[284], but not for Gessenia.

A problematic statement should encourage reflection about the assumptions that are held at that moment. A *problematic situation* is a situation in which a problematic statement is embedded, such as the following example:

> Charlie's math teacher says to her class: *"Every Wednesday you will have a math test"*. Imagine that today is Friday and his teacher applies a math test. Did the teacher lie to her class?

The teacher's claim is problematic in the sense that its mathematical interpretation is incompatible with the one usually given in everyday-language terms (compare with harmonious statements), where the given "every-statement" may be construed as an "if-and-only-if-statement". Thus, a problematic situation should challenge the participants to apply a new mathematical form of reasoning to a familiar statement, whose interpretation sometimes encourages a contrasting form of reasoning. This was particularly meaningful to Andrea during the 2018-intervention. As a result of the cognitive conflict that she experienced and as part of her reflection that "some" may mean "all", in contrast to her initial personal usage of it, she created two familiar statements in which she integrated her new mathematical assumptions (for details, see Chapter 5, Section II.1).

A discussion that includes a problematic statement has the main aim of breaking a pattern in the development of assumptions. It can support breaking the assumption that it is not possible to reason with false statements. For example, the statement *"All dogs are dinosaurs"* is problematic when its representation is debated, as happened with Gessenia and Lizbeth at the beginning of the second part of the 2018-intervention. Both teachers found it "weird" to represent a (false) statement merely based on its logical interpretation. They represented it as two separate sets, which is the actual relationship of the two sets involved in the statement. Similarly, Hawkins et al. (1984), who used *incongruent with practical knowledge* problems, suggest that making the participants aware that it is

[283] See Chapter 5, Section I.1.1.2 for details.
[284] See Chapter 5, Section II.1 for details.

hypothetically possible to reason with such premises, despite the arbitrary semantic relationships involved, can support the participants' attention to the logical structure.

The (false) statement *"For every natural number* n *(different from zero), when substituted in the formula* $1 + 1141n^2$*, the result is not a perfect square number"* is a problematic statement when discussing the status of confirming examples to prove USs due to the overwhelming number of confirming examples it has. It might create a conflict when determining the truth value of the statement as it is an unexpected false statement[285].

False USs that allow confirming examples were also problematic for Lizbeth as they granted access to her conflicting dual use of the term "justify" (for details, see Chapter 5, Section I.3.1.2). Barkai et al. (2002) also include a false universal statement that allowed confirming examples. Some elementary school teachers provided both confirming examples and counterexamples, but were unsure of the truth value of the statement. While in Barkai et al.'s study none of the teachers considered that their justifications would be accepted as mathematical proofs, Lizbeth was aware that a counterexample was conclusive to disprove a US. However, in other contexts, her personal use of "justify" might have been interpreted as if she was not aware of the sufficiency of a counterexample to disprove a US. It is possible that in a study similar to Barkai et al.'s, Lizbeth's response would have been counted in the group of teachers who incorrectly claimed that one confirming example justified that the statement was true, without further understanding the rationale behind her assumption. The case of Lizbeth, as well as Barkai et al.'s (2002) findings, show the importance of including false universal statements that allow confirming examples in proof-related discussions, where their truth value, justification and the validity of their justification are put forward. Lizbeth's case also underscores the difference between being aware of when confirming examples are sufficient or not to prove a universal-statement conjecture (which Lizbeth was) and being able to address generality in a mathematical proof (which Lizbeth was not). These are two different issues that encompass specific forms of reasoning and hence specialized background knowledge (e.g., methods and/or techniques like proof by induction, proof by cases, proof by contradiction, etc.) is needed for the latter (see Chapter 5, Section I.3.2).

Consider the general, but flexible, sequence of statements to be included in a discussion or sub-discussion (see Figure 61), which include five types of statements: familiar, abstract, imaginary, harmonious and problematic statements.

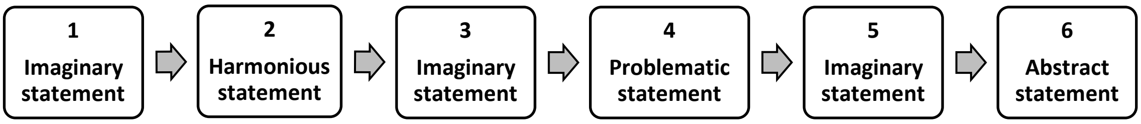


Figure 61. Organization of statements per discussion

The first statement I suggest using in a discussion is an imaginary statement (e.g., *"If a number is odd, it is «poponupulus»"*). This is particularly relevant in the case that the teachers' starting assumptions in relation to the main goal of the discussion have not yet been established (see Section 4.2 below). Using imaginary statements can facilitate the establishment and detection of absent or present *general* forms of reasoning that are held at certain moments (see the third and fifth type of statements suggested in Figure 61). Because imaginary statements are not framed in a specific context (e.g., mathematical or familiar contexts), they do not have specific truth values that might influence certain

[285] For details on how it supported Andrea's shift of her initial non-mathematical assumptions, see Chapter 5, Section I.3.1. For details on how it supported Lizbeth's exploration of different approaches in order to investigate the truth value of the statement, see Chapter 5, Section I.3.2.

forms of reasoning when used in tasks. The teachers' development of mathematical reasoning principles may be favored with the usage of imaginary statements. Their generic nature makes them a sort of bridge statement between concrete and abstract statements that might facilitate the establishment of general mathematical principles that should be formulated during the intervention. The use of imaginary statements is included to give access to the teachers' at-that-moment assumptions. For example, in relation to Discussion 2 and disproving USs, the UCS *"If a number is odd, then it is «poponupulus»"* could be proposed and a discussion about the characteristics for the mathematical evidence that would disprove it could be promoted. When an imaginary statement is used in a discussion, the unknown nature of the set(s) involved in it should be made explicit (e.g., by stating that "It is unknown what a *«poponupulus»* number is; however, that information is not crucial to solve the task").

The main purpose of following with a discussion of a harmonious statement is to elicit the teacher's related assumptions through the use of a statement whose common interpretation is not in conflict with the mathematical perspective. This may also facilitate the acceptability of a mathematical reasoning principle (as Epp, 2003, suggested). I advise suggesting harmonious statements that are compatible with the teachers' background knowledge, and reviewing terms in statements to ensure this compatibility. For instance, in the context of discussing the falsity of UCSs, the terms involved in the statement *"If a number is palindrome, then it is a square number"*, might have to be clarified to make it harmonious. It is possible that participants' assumptions might become evident when discussing harmonious statements that were not evident earlier. This is because of the specificity of the statement and might suggest that an assumption is not general.

As a next step, draw on imaginary statements to explicitly elicit the teachers' current "generic" assumption in relation to the mathematical principle in focus. Considering imaginary statements between harmonious and problematic statements allows verification of whether or not assumptions have changed at this stage. It might also support the generalization of the teacher's observations based on harmonious statements.

Problematic statements can be used to prompt the participants to call into question the assumptions they have established (when considering an imaginary statement). Including a problematic statement creates uncertainty. A problematic statement pushes the participant to think about what seems impossible or does not make sense to her/him so far. For example, the statement *"All numbers divisible by 6 are divisible by 3"* is problematic when discussing disproving USs, given that it is a true statement; however, a discussion about it promotes a conversation about "impossible counterexamples[286]" and it can later motivate an assumption about "proof" as eliminating counterexamples (see Chapter 5, Section I.3.3). It is demanding in the sense that it not only creates a dissonance in the teacher's body of assumptions, but it takes her/him a step forward to become aware of new forms of reasoning and apply them, which in a way were against her/his initial assumptions. Further work with imaginary statements can elicit again the teachers' current assumption(s).

Finally, an abstract statement can be used to establish generally the targeted mathematical reasoning principle. Resorting to abstract statements is one way to establish general forms of reasoning. For example, a discussion can finish with a focus on disproving USs by requesting the teachers to provide a characterization for all possible counterexamples that would disprove the abstract statement *"All X are Y"*. Similarly, before closing a discussion whose focus is on disproving ESs, you can ask for the characteristics of the

[286] For details on "impossible counterexamples", see Chapter 5, the introduction of Section I.2.

mathematical evidence that would disprove a statement of the form *"Some X are Y"* and whether it would be different from disproving the statement "*There exists X that is Y"* or "*There is X that is Y"*. As imaginary statements can also establish general forms of reasoning, the use of abstract statements may be omitted if the participants find them confusing.

4.2. Begin each discussion by identifying initial assumptions

The teachers' initial assumptions about a certain topic are directly linked to examples, non-examples, discussions, and in general their previous learning related experiences.

Finding out the teachers' initial assumptions supports making decisions on what assumptions need to be addressed during subsequent discussions, and what assumptions can be postponed, quickly discussed, or even overlooked.

It is important to become aware of the teachers' initial assumptions in relation to a targeted mathematical principle; however, it is crucial to be aware that the teacher might hold other related assumptions that can have a direct influence on it. During the 2018-intervention I learned that the teachers' assumptions were not isolated entities. I suggest aiming at consistency within the teacher's body of assumptions. In this sense it is crucial to find out and draw attention to the connections among the teachers' initial assumptions. When using the cognitive conflict approach to change a teacher's assumption, the expectation should be that other related assumptions are modified accordingly.

In the following sub-sections I refer to some considerations that should be made when identifying the participants' initial assumptions.

4.2.1. Be aware that interrelated initial assumptions can be accessed at different times in an intervention

The thing with interrelated initial assumptions is that sometimes the most fundamental assumption can be accessed later than the others and it may only allow a posteriori understanding of the development of the teachers' assumptions.

For example, during the exploratory interview of the 2018-intervention I learned about two of Gessenia's initial assumptions about the relation between a US and its converse[287]. I became aware of a third related assumption much later (In the UAS *"All X are Y"*, *X* and *Y* have the same elements [*X* = *Y*]). Gessenia's three initial assumptions were consistent with each other. Moreover, learning about Gessenia's third assumption helped me understand much more clearly why she used her first and second assumptions. It made sense that the most meaningful conflict Gessenia experienced during the development of her understanding of the relation between a US and its converse was directly linked to her third assumption. This suggests that her third assumption was more fundamental than the other two assumptions (for details, see Chapter 5, Section I.1.1.2).

Another example is the case of Andrea and her initial assumptions about "some-statements". They were all linked to each other and even related to the implicit negation of "all-statements" (*"Not all X are Y"* is equivalent to *"Some X are Y"*). Andrea reconsidered her body of assumptions as she accepted that a "some-statement" could be universally true, which was directly linked to the initial meaning she used for the

[287] These two assumptions are: If a UAS is (assumed to be) true, its converse is true (bAG1) and A UAS and its converse state the same (bAG2) (for details, see Chapter 5, Section I.1.1.2).

quantifier "some" (*"Some"* means *"from all, one group, but not all"*) (for details, see Chapter 5, Section II).

4.2.2. Use diagrams to elicit initial assumptions

Diagrams can be used to represent the logical interpretation of a statement. They grant access to the teachers' fundamental assumptions that can explain some other assumptions.

During the 2018-intervention the use of Venn or Euler diagrams to represent the logical interpretation of a statement was not compulsory. Yet, their limited usage made me aware of some of the teachers' assumptions that were not aligned with the mathematical perspective. For instance, although Andrea was aware that a universal statement could be represented by the *set of analysis* included in the *conclusion set*, Andrea assumed that this relation necessarily implied that the set outside the set of analysis, but still in the conclusion set, was non-empty (see Chapter 5, Section I.1.1.1 for details). In contrast, Gessenia's attention to a diagram for a UAS made me aware of the grounds for her assumption that a UAS implied its converse. Gessenia initially assumed that both sets involved in a UAS had exactly the same elements (see Chapter 5, Section I.1.1.2 for details). Similarly, even though the representation that Andrea suggested for the logical interpretation of "no-statements" was as expected (two disjoint sets), she assumed that the "no-statements" referred to the elements outside the set of analysis (see Chapter 5, Section III.1.1 for details).

4.2.3. Be aware that some assumptions may stem from the intervention

During the 2018-intervention I realized that it was important to include a discussion to find out the initial assumptions the teachers used to disprove USs. This was particularly relevant as I noticed some other related assumptions that Andrea and Gessenia had made in previous discussions. On one hand, Andrea assumed that a counterexample was not sufficient because she considered it a concrete example, which she had overgeneralized was never a valid justification (for details, see Chapter 5, Section I.2.1.2). On the other hand, Gessenia assumed that the counterexample was not acceptable because of its form of expression; the arguments we had accepted as valid justifications earlier had always had a text or narrative form (for details, see Chapter 5, Section I.2.1.1). Because of that discussion, I am now more aware of the possible assumptions that might emerge during a similar intervention and the importance of addressing them.

I suggest beginning Discussion 4 (about proving USs) by asking the teachers to provide an example of a true US and to explain how they know that it is universal and true. Their answers can reveal not only the teachers' criteria for regarding a US as a true statement, but also whether the teachers spontaneously give a US that applies to a finite number of cases or not. In previous interventions I observed that teachers tend to overgeneralize that USs always apply to an infinite number of cases (see e.g., Chapter 5, Section I.3.1.1). This might be due to their previous experiences with USs that involve infinite cases, including their experiences during the intervention.

4.2.4. Consider possible reorganizations of the discussions

Learning about the teachers' initial assumptions may suggest a rearrangement of the discussions for an intervention. The decision on omitting, including or changing the order of the discussions can be made based on what initial assumptions are identified. The teachers who participate in future interventions might not necessarily hold the same initial assumptions as the teachers in the 2018-intervention. Being open to reorganizing an intervention can support a better development of the teachers' assumptions. The important thing is that there is a need to be clear of what assumptions are targeted in each discussion.

4.2.5. Use imaginary statements

A task that involves an imaginary statement (e.g., *"Each «bogui» is a «fantaslopitocus»"*) does not demand awareness of practical or mathematical content knowledge to be solved (for details on *imaginary* statements, see Chapter 4, Section II.2.1, Stage 1.3). For example, during the exploratory interview of the 2018-intervention I asked the teachers to find statements that made the same claim as the imaginary statement *"All «Vallejo» numbers are even numbers"*. The teachers were told that it was irrelevant what "Vallejo" numbers were in order to solve the task. The teachers' responses gave me access to different initial assumptions that the teachers held. On one hand, Gessenia and Lizbeth assumed that a US and its converse made the same claim; while Andrea assumed that given a true US, its converse was false (see Chapter 5, Section I.1.1). Their initial assumptions were also corroborated later, during the discussions.

Because imaginary statements have an absence of content, they facilitate access to general assumptions or forms of reasoning that individuals hold about single-quantified statements at a certain moment.

Some of the teachers' initial assumptions can be identified before the discussions take place, as I did during the exploratory interview of the 2018-intervention; otherwise, they can be determined right at the beginning of each discussion. In any case, using imaginary statements supports that goal as it makes those general assumptions or forms of reasoning accessible without the influence of the truth value of the statements or familiarity with the content.

4.3. Finish each discussion explicitly establishing the new forms of reasoning

Establishing new forms of reasoning in an explicit way is as important as finding out what the teachers' initial assumptions are. As I suggested above, the use of *abstract statements* or *imaginary statements* might facilitate this stage (see Section 4.1). The main goal is to make sure that the teachers can articulate and explain the general mathematical reasoning principles targeted during each discussion, in case they have not been previously come up with them.

There are at least two ways to keep an explicit record of those new forms of reasoning. One is that the teachers can use a sort of personal diary where they can write down their insights about the discussions they engaged in during the intervention (see Section 7.5.7 below for more details). Another way is sharing a public space where everybody posts their ideas and helps to refine new forms of reasoning as a community. For example, during the teachers' lessons they noted the theorems their students came up with in front of the class and asked them for refinements. Similarly, this is what we (the teachers and I) did with the theorems the teachers proved in relation to the mathematical content during the first part of the intervention, though we did not follow a similar approach with the targeted reasoning principles during the second part of the intervention. Following this approach revealed that some teachers struggled to lead their classes as they themselves had a hard time expressing their ideas in a clear way. So, having the teachers take notes of principles, as they did with their students, may be helpful. Inviting the teachers to engage in this work might indirectly support an enhancement in the way they express their ideas as they need to "negotiate" with other members of the community to agree on clear ways to formulate those principles. The teachers should agree on the principles they identify as well as on their explanations, as members of a community. This would make them aware of the principles they agree on and the progress they make together.

The reasoning principles should be easy to visualize later when other discussions take place. Big sheets of paper on the walls can be used for this. The principles should be named as they arise and expressed in the teachers' own wording. There should be an expectation that those principles could later be refined or modified according to new observations that might emerge. If modifications are to be made, include expressions like "version 1", "version 2", to distinguish the principles and track the way they evolved. This might also be a good way to inform future changes to the design of the intervention.

5. Principles about language

"As Pimm points out, although people talk about how precise mathematics language is, the precision actually is not in the language itself, but in the way it is used."

(Schleppegrell, 2007, p. 150)

"But 'glory' doesn't mean 'a nice knock-down argument'," Alice objected.

"When I use a word," Humpty Dumpty said, in rather a scornful tone,
"it means just what I choose it to mean—neither more nor less."

(Carroll, 2009, p. 190)

The way individuals interpret proof-related terms and expressions will undoubtedly influence their reasoning when engaged in dis/proving. The way those words are used in everyday language can have a direct impact on the way individuals dis/prove in mathematics. There are several examples of terms and expressions that are used in everyday life with a certain connotation but when they are used in the *mathematics register* (Halliday, 1978), their meaning changes. Examples include "some", "none", "there is" and "proof". Nevertheless, as Lee and Smith (2009, p. 24) note, "in the studies of proof schemes and proving, the effects of the students' conception of these mathematical foundations are usually ignored".

In general, when these terms or expressions are used in the formulation of mathematical statements, the influence of the everyday-context interpretation can jeopardize the understanding of their logical structure and therefore interfere with other proof-related activities. The linguistic challenges of the mathematics register suggest paying close attention to language and involves, as Schleppegrell (2007) put it, "focusing on the features of the language through which mathematics is constructed" (p. 147). The differences in the meaning of quantifiers and connectives can affect the way individuals argue about the truth value of mathematical statements, depending on the interpretation used for those quantifiers and connectives. I concur with Lee and Smith (2009) that "students' learning to prove is likely complicated by the cognitive challenge of interpreting the proving task in accordance to its logical meaning instead of its pragmatic meaning" (p. 23).

In particular, I observed this influence in the teachers' proof-related assumptions during the 2018-intervention. For example, language had an impact in the development of Andrea's assumptions about "some-statements" (for details, see Chapter 5, Section II.1), the teachers' assumptions about what constitutes a proof (see e.g., Chapter 5, Section I.3) and about negation (see e.g., Chapter 5, Section II.3).

Lee and Smith (2009) provided the following task where the interpretation of the quantifier "some" is involved and different register interpretations might come into play.

"All A are B. Does it follow that some A are B?"

The authors anticipated that the answer would be "yes" if it were interpreted mathematically, but "no" given a pragmatic (everyday) interpretation.

When everyday language is used to interpret mathematical quantifiers and connectives, it can hinder communication. It is important to identify particular patterns in the relationship between everyday and mathematical language to support students' development of a more sophisticated mathematical language (Hempel & Buchbinder, 2022).

5.1. Pay attention to the terms used in proof-related discussions and their interpretation

Supporting the development of teachers' proving skills should involve careful attention to the way that teachers use certain terms (e.g., "proof/justification") as well as the way this evolves during their development. It is possible that the use of some terms differs among the participants. For instance, during the 2018-intervention the teachers tried to represent the statement *"All numbers divisible by 6 are divisible by 3"*. The teachers engaged in a heated discussion as they disagreed strongly about their explanations for what the representation of the statement should be. The source of their disagreement was the divergent interpretations they used for "universal set". The expression was introduced by Gessenia, who claimed that the "universal set" in the statement was the set of numbers divisible by 3, which Andrea agreed with. In contrast, Lizbeth claimed that the "universal set" was the set of numbers divisible by 6. Although the three teachers reached the same representation for the statement (i.e., *the set of numbers divisible by 6 included in the set of numbers divisible by 3* shown in an Euler diagram), their interpretations were different and the conflict stemmed from this difference. While Andrea and Gessenia interpreted "universal set" as an *encompassing set* or *the set that included the set of analysis*, Lizbeth interpreted it as *the set of analysis, the set for the search* or *the set to which the universal quantifier applied to*. Becoming aware of the interpretations of key terms and expressions that participants use, and explicitly exposing them, are crucial activities that should be part of all (and in particular proof-related) discussions. Lack of awareness of possible mismatches in the use of terms and expressions might lead to endless chaotic discussions.

Lizbeth's use of "justification" during the 2018-intervention also shows the importance of identifying the teachers' initial use of terminology and paying close attention to the way it might change. It shows the significance of paying careful attention to the teachers' utterances when proof-related discussions are fostered, so that a real sense of how their assumptions develop can be attained; otherwise, hasty conclusions can be drawn. For example, at first glance the case of Lizbeth might suggest that during the intervention her judgements of arguments for infinite universal statements were based on an empirical proof scheme. However, her interpretation of "justification" at that time suggests otherwise (for details, see Chapter 5, Section I.3.1.2).

This might also be the case for terms that are important when having discussions related to proof. For instance, Andrea's use of the term "logical" was closer to common-sense than it was to a logical perspective (see Chapter 5, Section II.1); Andrea's use of "not necessarily" was far from its mathematical meaning (see Chapter 5, Section I.1.1.1, Episode 2); and Andrea's initial use of "verify" to mean test with the use of examples was also based on everyday language (see Chapter 5, Section 2.1.1, Episode 4).

6. Principles for developing meanings

Among these principles I include providing initial incomplete inputs for the meaning of key terms and requesting refinements (Section 6.1) and labeling some important terms and expressions to avoid confusion (Section 6.2).

6.1. Provide initial incomplete inputs for the meaning of key terms and request refinements

In general, individuals hold personal meanings for the words they use. Explicit agreement on the way we use key terms is crucial when engaged in proof-related discussions. This is true not only at the level of the community of mathematics educators who publish articles about proof, argumentation, justification and so on (G. J. Stylianides, A. J. Stylianides & Weber, 2017; Balacheff, 2002), but also for teachers in training programs, or school students in classroom contexts, to engage in productive discussions.

Baggini and Fosl (2010) point out that,

> A clear definition of how you will use the term thereby both helps and constrains discussion. It helps discussion because it gives a determinate and non-ambiguous meaning to the term. It limits discussion because it means that whatever you conclude does not necessarily apply to other uses of the term. As it turns out, much disagreement in life results from the disagreeing parties, without their realizing it, meaning different things by their terms. (p. 32)

Even though I agree for the most part with Baggini and Fosl, in my view there are some concepts for which it may not be possible to provide a thorough definition that combines all relevant aspects at once, before engaged in related discussions. For instance, key concepts such as “proof” in mathematics may not be as simple to define in comparison to mathematical concepts such as “continuous function” or “matrix”. It is not easy to encompass all the important aspects of “proof” in a definition, especially when individuals do not have a mathematical background.

There are a range of ways participants might come to understand what “proof” means in mathematics: We could provide a definition of “proof” and expect that teachers determine whether an argument satisfies all its conditions. We could instead show them examples and non-examples of “proofs” and expect them to identify themselves what makes those arguments a “proof”. We could guide them towards their own realization of what makes an argument a “proof”. Or we could simply teach them methods of proof and instruct them to apply those methods.

In my own experience I learned about proof[288] by proving. This is similar to the process described in Davis and Hersh (1998, pp. 34–44). No professor defined what a “proof” was for me. Nonetheless, we were doing proofs and learning sound forms of mathematical reasoning through proving. I learned methods of proof in that way. Nobody explained to me why those methods were “allowed” or “worked” in mathematics. No professor explained to me, for example, why proving the contrapositive of a conditional statement resulted in proving the conditional statement itself. Truth-tables supported my understanding at some level; however, truth-tables did not give me a real insight into why those methods worked the way explanations can.

[288] I studied mathematics in the university.

Deductive reasoning is the conclusive mathematical mode of argumentation; however, teaching deductive reasoning to primary school teachers or students in an abstract way might not be as meaningful as having them learn examples and non-examples of this form of reasoning. From the 2018-intervention I concluded three things related to this: First, natural language has a big influence on the way teachers interpret and reason with mathematical SQ-statements (see Section 5 above). Second, there might be alternative ways to have fruitful discussions about proof-related issues without completely defining "proof" or related concepts beforehand. Third, a conceptualization of proof can be gradually developed as a *cluster concept* or *cluster category* (see Weber, 2014; Czocher & Weber, 2020).

There are key concepts in mathematics that can be gradually conceptualized without the need to provide a complete definition initially. I found that to be the case for "justification", "counterexamples" and the quantifier "some". The teachers actively engaged in discussions about their most important features. I believe that some concepts might have a bigger impact on the development of teachers' understandings when their meanings evolve rather than being completely predefined.

Additionally, the way key proof-related terms are used is heavily influenced by the everyday-language contexts in which they are probably already used and where their interpretations are sometimes distant from the mathematical context. This certainly adds more challenges to the development of assumptions about those terms. In this regard, Schleppegrell (2007) claims that "[l]earning the new vocabulary that is centrally mathematical may be easier than learning the technical meanings for words that students already know in other contexts." (p. 142). Schleppegrell also notes that,

> a key challenge in mathematics teaching is to help students move from everyday, informal ways of construing knowledge into the technical and academic ways that are necessary for disciplinary learning in all subjects… Learning the language of a new discipline is a part of learning the new discipline; in fact, the language and learning cannot be separated. (p. 140)

This suggests an approach that allows a transition from the natural language register to the mathematical register (Halliday, 1978, p. 195), as I show in Figure 62.

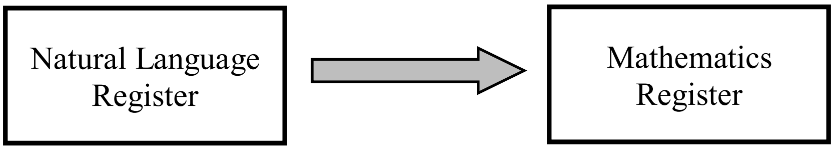


Figure 62. The transit of the assumptions related to key proof-related terms that are also used in everyday life.

In this context, I recommend that a new intervention provides initial inputs that are consistent with mathematical interpretations but not in conflict with natural language. The participants should take those inputs as "starting-point meanings", which they will be expected to progressively refine and make more accurate by engaging with conflicting cases (see Figure 63). I am not suggesting introducing new mathematical terms and expecting that the teachers complete the gaps in their conceptualization on their own, without a plan. My suggestion is that, if a concept is introduced using this general approach with the expectation to be "refined" by the teachers, then relevant activities that support those refinements should be planned in such a way that they facilitate those conceptualizations and attention is not distracted from this goal.

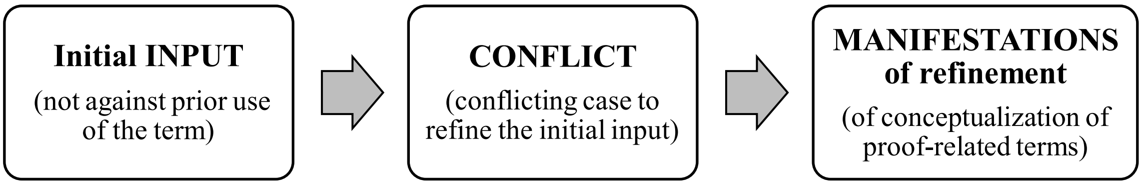


Figure 63. Suggestion for the path towards a conceptualization of key proof-related terms

In what follows I include suggestions for ways of conceptualizing three important proof-related terms: "justification", "counterexample", and "some". While the first and third terms require a shift of register as the teachers use them in their everyday life, the second is usually a new technical term for the teachers.

6.1.1. The terms "justification" and "justify"

From the moment the word "justify" is included in a discussion, there is a need to agree on an initial shared meaning for it even if it is not completely accurate from a mathematical perspective.

As I suggested above (see Figure 63), an intervention should provide an initial input. That input should not be in conflict with prior usage (e.g., everyday context) but should also be consistent with its mathematical meaning. The initial input I suggest for "justify" allows the teachers to reflect on what mathematical reasons can explain why[289].

> *Justify X* means explain why *X*; provide the reasons for *X*.

Moreover, an explicit request to refine the meaning of "justification" directs the teachers' attention to the aspects that make a justification a mathematically valid justification; for example, a slide can include a reminder such as:

> Henceforth we must focus on refining the concept *justification* from a mathematical perspective (i.e., a mathematically valid justification).

The whole intervention can be seen as aiming at shaping the conceptualization of "justification". That is because justifications might look different according to the statement under discussion and the different types of statements discussed during the intervention.

Including incomplete inputs for terms such as "justification" was particularly relevant during the 2018-intervention because of the different ways that teachers conceptualized it. For example, Andrea's conceptualization of *justification of true USs* included her division of USs into categories where her focus was on the number of cases the USs applied to (see Chapter 5, Section I.3.1.1). On the other hand, Lizbeth's conceptualization of *justification of a true US* involved her insight that proving USs involves the non-existence of counterexamples (see Chapter 5, Section I.3.3) and also includes her integration of and differentiation from new terminology (e.g., "sufficient to guarantee"), depending on the status of confirming examples (see Chapter 5, Section I.3.1.2).

---

[289] My use of "explain" herein should be seen in broad terms and is not necessarily intended to refer to the explanatory function that proofs might have (see e.g., Hanna, 2000, 2018).

6.1.2. The term "counterexample"

Based on my 2018-intervention, it is likely that primary school teachers will not be familiar with the term "counterexample", at least not in the mathematical context[290].

I suggest that Discussion 2 include an explicit written initial prompt to use as a reference for "counterexample", such as this one:

> A *counterexample for the statement S* is an example that contradicts *S*; an example that shows that *S* is false.

This prompt should be given after the teachers are asked to find out the truth value of a (false) universal statement, a sufficient justification for it (a counterexample) has been identified, and an explanation has been put forward that shows why the counterexample is a valid mathematical justification. Subsequent discussions should aim at supporting the teachers' refinement of their conceptualization for counterexample (see Section 7.3 below for concrete suggestions on how to refine this concept).

6.1.3. The term "some"

The teachers began the 2018-intervention with an initial assumption for the meaning of the quantifier *some*. That meaning relied on the everyday register, where "some" is usually interpreted as *some, but not all* (Epp, 1999).

The term "some" was a very challenging case since its meaning according to the everyday context already triggered different sorts of non-mathematical assumptions in relation to "some-statements". For instance, if it is true that *"Some X are Y"*, then this implies that *"Some X are not Y"* (see Chapter 5, Section II.1 for details). Similarly, if *"All X are Y"* is true, then *"Some X are Y"* is false (see Chapter 5, Section II.2) and the negation of *"All X are Y"* is *"Some X are Y"* (for details, see Chapter 5, Section I.4).

In this case the initial input for the meaning of the term "some" can be given, as *at least one*. This input was taken up by the teachers without any opposition during the 2018-intervention. However, Andrea only reconsidered her initial personal use of "some" when a conflicting case (a true existential affirmative statement that is universally true) was put forward (see Section 7.1.2 below for a specific suggestion on how to refine the initial input for "some").

Without a conceptualization of these three terms aligned to mathematics, the understanding of dis/proving SQ-statements cannot be complete from a mathematical perspective.

6.2. Label some important terms and expressions to avoid confusion

Some terms might not require a sequence of refinements like those in Section 6.1. These expressions are simple enough to define and understand. They can be seen as foundational to the development of other understandings, including of the terms in Section 6.1. For example, expressions like "set of analysis" and "conclusion set" support the understanding of the logical interpretation of a statement; expressions like "confirming examples" and "irrelevant examples" support the development of a characterization of counterexamples.

[290] In the introduction of Section I.2.2 in Chapter 5, I explain that Gessenia and Lizbeth did not use the term "counterexample", while Andrea used it, but with a meaning from a context different from proving statements.

Those definitions should be given (perhaps written on slides or a handout) as soon as they first arise. In this way the teachers have access to them as they need them. Figure 64 is an illustration of how this can be done.

Discussion 1A

Consider statement S1.

S1: All divisions of natural numbers are exact divisions

Now answer the following questions:

a) What are the two sets involved in S1?

b) From those two sets, which one does S1 refer to?

INPUT[Def]: Set of analysis in a statement S – the domain set in S; the set S applies to; (in Copi et al.'s [2014] terms the "subject class")

INPUT[Def]: Conclusion set in S – the set which description is given by the conclusion condition; (in Copi et al.'s [2014] terms the "predicate class")

c) Provide three examples of elements from the *set of analysis*.

d) What elements from the set of analysis does S1 refer to? (all of them? Some of them?)

e) What does S1 assert about those elements?

f) Represent S1 in such a way that its representation expresses exactly what S1 states. Use a Venn or Euler diagram.

Figure 64. Definitions of "set of analysis" and "conclusion set" on a slide.

Teachers need explicit written definitions (marked "INPUT[def]" in Figure 64) so that they have access to them as often as they want. This can support their reflections and can show them that definitions in mathematics are important, in particular when proof-related discussions take place.

This suggestion is especially relevant in the case of the term "example" and the different types of examples. I suggest using different expressions for the subcategories of examples that are used in the intervention: "confirming examples", "irrelevant examples" and "counterexamples" (see Chapter 3, Section I.2.2). Do not call them all "examples". This was a limitation of the 2018-intervention as it became difficult sometimes to interpret the teachers' claims about "examples" during the intervention. For example, at the beginning of Discussion 1.0 the agreement was that the term "example" was used to mean a "confirming example"; however, Andrea[291] sometimes used it to mean "contradicting example", and in the following case to mean specific cases or calculations that belong to the set of analysis of the statement (i.e., "relevant examples").

> Andrea: *Besides, they [the students] don't know if it [the conjecture] is true or false. So, the examples will help them, BUT as a justification, ONLY for false universal statements.*

Similarly, during her teaching Gessenia sometimes used the term "example" to mean "a specific case" or "counterexample", in different contexts. In the following extract of her teaching of Session 12 she means "counterexample":

> Gessenia: *He (Julito) says that always and we have seen that not always. An example, an example that breaks the claim he made is enough.*

The teachers' inconsistent use of "example" might have been confusing to their students. Having and using precise terminology in cases like this will help teachers and also their students to provide more precise conjectures, explanations and assertions.

[291] For details on the context of this episode, see Chapter 5, Section I.3.1.1.

7. Principles about techniques

In this section I suggest five techniques that I have found helpful to guide the development of teachers' thinking about proof-related issues. They include using cognitive conflicts to change assumptions (Section 7.1), leveraging others' arguments to introduce valid and invalid forms of reasoning and submit them to criticism (Section 7.2), using irrelevant and confirming examples to refine the characterization of counterexamples (Section 7.3), using indirect approaches to show the truth of universal statements and falsity of existential statements (Section 7.4), and paying attention to signs of emerging understandings to follow the course of the development of the assumptions (Section 7.5).

7.1. Use cognitive conflicts to change assumptions

I suggest using cognitive conflicts to shift some important proof-related assumptions. Cognitive conflicts can make teachers aware that their assumptions do not always hold by creating dissonance between their assumptions. I used this approach during the 2018-intervention and successful conflicts supported the teachers' reconsiderations of some of their initial assumptions.

Limón (2001) points important steps that should be considered in the cognitive conflict approach[292]. Here I summarize them in Figure 65. First a confirmatory example (i.e., an example that is aligned with the initial assumption) is presented. Its purpose is to draw awareness to the initial assumption. Second, a conflicting example (i.e., an example that is in conflict with the initial assumption) is introduced. Because attention was first drawn to the initial assumption by the confirmatory example, it should be easier to notice a contrast between the initial assumption and new observations based on the conflicting example. A discussion that includes explicitly addressing those differences is needed. Finally, a new (or the previous) confirmatory example is used as an object for reflection, in order to better understand the conflict and its resolution.

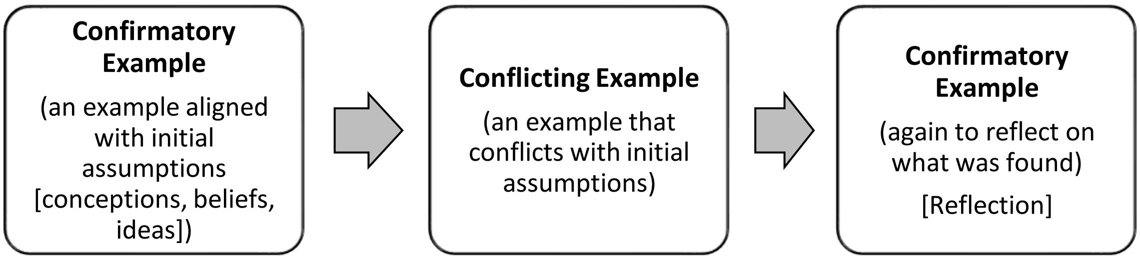


Figure 65. The main structure of the cognitive conflict approach

Observe that among the types of statements that I suggested using, *problematic* statements can be seen as those that are expected to trigger cognitive conflicts (see Section 4.1 above).

In some circumstances where the cognitive conflict approach is used, an individual might not recognize the conflict (Tirosh & Graeber, 1990). In that situation, working with a group of teachers instead of one teacher at a time is advantageous. In a group, there might be other teachers who perceive the conflict and be able to bring awareness of it to the discussion.

One reason a conflict might not be recognized is lack of background knowledge (Gal, 2019). This is why I suggest beginning proof-related interventions with a first part with a focus on mathematical content knowledge. Establishing the mathematical concepts to be

[292] See Chapter 3, Section I.1.1, for details.

used during later proof-related discussions should ensure that conflicts are not overlooked due lack of background knowledge.

In the following sub-sections I review five initial assumptions related to proof that the teachers held during the 2018-intervention and the way cognitive conflicts played an important role in their reconsideration. Many of these assumptions were adopted based on one or few instances. For each I suggest conflicting examples that might create conflict and support changes. In the 2018-intervention I could not evaluate whether all these initial assumptions remained changed after the intervention. Ideally, such an intervention should investigate ways to determine whether or not new assumptions persist.

7.1.1. An infinite universal statement can be proved with examples

Verifying a number of confirming examples is a common mode of argumentation that individuals use when requested to prove a generalization that involves infinite cases (e.g., Buchbinder & Zaslavsky, 2007; Chazan, 1993; Healy & Hoyles, 2000; Knuth, Chopin, Slaughter & Sutherland, 2002; Morris, 2007; Sowder & Harel, 2003). It has been usually claimed that students who produce that type of arguments use an empirical scheme of justification (e.g., see Harel & Sowder, 1998, 2007; G. J. Stylianides & A. J. Stylianides, 2009).

My work has shown that difficulties around examples extend beyond an assumption that sufficient examples can prove a universal statement. In that respect, I concur with Lockwood et al. (2016) that:

> [r]ather than simply emphasizing that examples do not constitute proof, we instead need to help students learn ways to build bridges between the conviction and understanding that can come with examples and deductive arguments (p. 193).

There are certain statements that are exceptions to the general rule that examples do not constitute proof. An exhaustive set of confirming examples is sufficient mathematical evidence to prove universal statements with a "small" finite number of cases involved (e.g., *"All one-digit numbers divisible by 6 are divisible by 3"*). Theoretically all finite USs can be proved by exhaustion if every case involved is verified, though this is not practical when large numbers of cases are involved. Existential statements (e.g., *"Some prime numbers are odd numbers"*) can be proved with a single example. Furthermore, there are situations where examples may qualify as generic examples in valid proofs for true universal statements.

In Figure 66 I include the assumptions that the teachers might hold in a sequence of discussions using cognitive conflicts to refine the assumption that examples cannot prove.

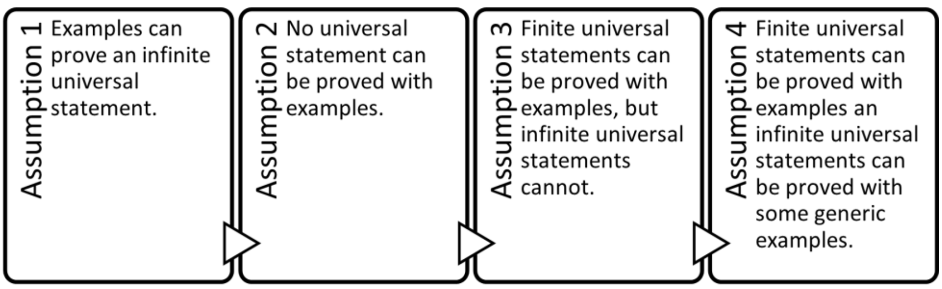


Figure 66. The possible development of the teachers' assumptions about proving universal statements that involve using the cognitive conflict approach

Assumption 1: Examples can prove an infinite universal statement

Making the teachers aware that verification of a few confirming examples does not prove infinite universal statements can be achieved with the support of two conflicting statements:

Conflicting Example 1[293]: *All palindrome numbers are divisible by 11*

Conflicting Example 2: *For every n natural number (different from zero), the expression:* $1 + 1141n^2$ *does not produce a perfect square*

Both conflicting statements are false universal statements that admit confirming examples. Both involve a request to determine the truth value of the statement as well as a justification for it. These confirmatory examples can be used in a discussion both to confirm Assumption 1 as well as to raise a conflict with it. For example, Conflicting Example 1 can be included in a discussion slide (see Figure 67).

Discussion 3B

Peter has formulated the following conjecture:

St2: All palindrome numbers are divisible by 11

(INPUT[Def]: A **palindrome number** is the number that reads the same backwards)

Peter claims that his conjecture St2 is true because, as he argues,

"I have verified it with large numbers. 121, 2442, 3553, 123321, they all are divisible by 11"

He used his calculator to support his work.

Now answer:

a) Is Peter's conjecture (St2) true as he claims? Why?
b) Is Peter's justification (St2) a mathematically valid justification? Why?
c) What is the status of Peter's conjecture then?

Figure 67. Slide suggested for Discussion 3B

The confirmation of Assumption 1 comes from Peter's argument. The way it is formulated reinforces the teachers' initial assumption that a few confirming examples could prove an infinite universal statement. The teachers' acceptance of Conflicting Example 1 as true signals their use of Assumption 1. The conflict with Assumption 1 emerges as they realize that Conflicting Example 1 is false and suggest their own counterexamples to disprove it. I used this discussion during the 2018-intervention and the teachers easily found out that the conjecture was false. If the teachers do not realize that the conjecture is false, the teachers can be asked whether the conjecture is *always* true, and if this does not prompt them to think of a counterexample, they can be asked if number 101 is a palindrome number[294] and can be divided by 11.

The Conflicting Example 2 is included in the "Monstrous Counterexample Problem" that G. J. Stylianides and A. J. Stylianides (2009) use in their study with pre-service teachers. The statement is false, but it is very hard (if not impossible) to disprove with a hand calculator or pencil and paper. The smallest counterexample to it is a number of 26 digits.

The Conflicting Example 2 admits an overwhelming number of confirming examples while involving mathematical concepts that are basic enough to be understood by primary

[293] An alternative contradictory example for Contradictory Example 1 is the one included in an activity the teachers solved at the beginning of discussion 3 of the 2018-intervention (see Appendix A12).
[294] I include a definition of a palindrome number in the slide to avoid distractions from the main aim of the discussion.

school teachers. For the 2018-intervention, I adapted the Stylianides and Stylianides' version of the problem[295]. Instead of presenting it as a given statement and providing facts about it, I asked the teachers to try out some examples, identify the pattern and formulate a conjecture, which would be Conflicting Example 2[296]. The main goal of presenting the discussion in this way is that the teachers could increase their confidence that the Conflicting Example 2 is true and they could attribute that confidence to the many confirming examples the statement has. This would play a confirmatory role to their initial assumption (Assumption 1) and test whether they resorted it again. The conflict is expected to emerge as they realize that there is a counterexample for the statement, which they did not expect[297].In the 2018-intervention, Conflicting Example 1 already supported the teachers' rejection of Assumption 1. However, the teachers might have assumed that the statement was false only if they could easily find a counterexample. To investigate whether rejecting Assumption 1 depends on being able to find a counterexample, I suggest using Conflicting Example 2 afterwards. This supports the teachers' forming a new assumption that even though there might be strong empirical evidence for an infinite US, it might have unexpected or very hard to find counterexamples.

As was the case in the 2018-intervention, after discussing these two conflicting examples the teachers might shift to a more extreme assumption: *No universal-statement conjectures can be proved with examples* (Assumption 2).

Assumption 2: No universal statement can be proved with examples

The teachers' realization that there might be cases of universal statements that can be proved with examples can be supported with the following conflicting statement:

Conflicting Example 3[298]: *All natural numbers smaller than 5 are smaller than 7*

Observe that this is a statement that applies to a finite number of cases. A discussion in which the teachers are asked for the number of cases involved in Conflicting Example 3 and what would be needed to justify that it is true can foster the change of Assumption 2 to a new assumption: *Finite universal statements can be proved with examples, but infinite universal statements cannot* (Assumption 3).

Assumption 3: Finite universal statements can be proved with examples, but infinite universal statements cannot.

The teachers' awareness that there might be cases of examples that qualify as proof can be fostered with discussions about generic examples.

Illustrating the features that make a generic argument a proof can be more insightful than just telling the teachers what characteristics a generic proof should have. Include both generic arguments that qualify and do not qualify as proofs, and, more importantly, discussing why in each case (see Reid & Vallejo-Vargas, 2018).

[295] They also report on the same problem in A. J. Stylianides & G. J. Stylianides (2014) and G. J. Stylianides and A. J. Stylianides (2014).
[296] For details on a discussion where the Conflicting Example 2 was expected to emerge, see Appendix D3, and focus on the Task and Discussion (Task), after Discussion 3.2.
[297] Like G. J. Stylianides and A. J. Stylianides (2009) did, I included the counterexample in a slide (see Discussion (Task), after Discussion 3.2 in Appendix D3.
[298] This statement was included in Discussion 4 of the 2018-intervention (see Appendix D4). An alternative contradictory statement could be *"All one-digit natural numbers divisible by 6 are divisible by 3"*.

As I have shown in my work, moving from using examples to verify to relying only of proofs involves the formation of assumptions that may not be mathematically accurate and a discussion about them should be included.

7.1.2. "Some X are Y" implies that "Some X are not Y"

The teachers in the 2018-intervention used their everyday meaning of "some" ("some, but not all") in a mathematical context (see Chapter 5, Section II.1 for details). As a direct implication of this assumption, they concluded that *if some X are Y, then some X are not Y* (Assumption 4). In this section I outline how to use a cognitive conflict to trigger change of this assumption.

The confirmation of Assumption 4 can be elicited using statements St1 and St2.

> St1: *Some divisions of natural numbers are exact divisions*
>
> St2: *Some divisions of natural numbers are not exact divisions*

Observe that St1 is a true affirmative "some-statement" and St2 is a true statement, which is the negative "some-statement" of St1. It can be supported by drawing the teachers' attention to the truth of St1 first, and then asking them whether this guarantees that St2 is also true.

It is possible to make the assumption (*"Some X are Y" implies "Some X are not Y"*) explicit by introducing it as someone else's claim. For example,

> Kevin has made the following claim about St1 and St2: "I do not need to prove that St2 is true. It is true because St1 is true. If some divisions are exact, then some are not."

Asking the teachers whether they agree with Kevin's claim and to explain their answer can elicit their Assumption 4. Reminding the teachers the initial input for the meaning of "some" is *at least one* (see Section 6.1.3 above) might not change the truth value they ascribe to the statements; however, it might begin to trigger hesitation on whether the truth of St1 implies the truth of St2.

The teachers' change of Assumption 4 can be achieved by using statements St3 and St4.

> St3: *Some numbers divisible by 4 are even numbers*
>
> St4: *Some numbers divisible by 4 are not even numbers*

Observe that St3 is a true affirmative "some-statement" and St4, its respective negative "some-statement", is false. Both make the conflicting example that will support changing Assumption 4. Their realization of the conflict should be achieved by drawing the teachers' attention to the truth value of the statements and asking them whether the truth of St3 implies the truth of St4. As with St1 and St2, both statements are mathematically contextualized, but St3 is universally true, which implies that St4 is false. This conflicts with the teachers' Assumption 4, which might lead to a reformulation of the assumption, as was the case of Andrea during the 2018-intervention[299].

7.1.3. Given a true UAS, it follows that its converse is true (or false)

Considering that a true UAS implies that its converse is necessarily true is a common assumption that individuals use (e.g., Epp, 2003; Hoyles & Küchemann, 2002; Yu, Chin & Lin, 2004, for details see Chapter 2, Section I.1). During the 2018-intervention, Gessenia and Lizbeth used this assumption (see Chapter 5, Section I.1.1.2); whereas

[299] For details, see Chapter 5, Section II.1.

Andrea assumed that given a true UAS, its converse was false (see Chapter 5, Section I.1.1.1). Andrea's assumption is less common, but it can emerge, possibly as an overgeneralization arising from awareness that the more common assumption, a true UAS implies that its converse is true, is false. I suggest discussing both assumptions in one discussion. If Andrea's assumption arises, it can be addressed at the same time. Consider the assumptions as follows:

> Assumption 5: Given a true UAS, it follows that its converse is true
>
> Assumption 6: Given a true UAS, it follows that its converse is false

The same confirmatory example can be used for both Assumption 5 and Assumption 6, using imaginary statements. In this way other variables such as the content or the truth value do not interfere with the teachers' consideration of their initial assumptions. For example, use St5 and St6.

> St5: *If a number is even, then it is «bogui» (or All even numbers are «boguis» numbers)*
>
> St6: *If it is a «bogui» number, then it is even (or All «boguis» numbers are even numbers)*

Ask whether St5 implies St6 and whether they make the same claim. Also ask for a representation of St5 and whether in the representation one can be sure that there are elements that are *«bogui»* and not even. These requests are important as the 2018-intervention showed due to their close relationship[300].

A conflicting example for Assumption 5 is given first. A true familiar UAS and its false converse are the conflicting example. Using statements whose truth value is within the background knowledge of the teachers, such as St7 and St8, might support their realization of the conflict.

> St7: *All human beings are mortals (or If it is a human being, then it is a mortal)*
>
> St8: *All mortals are human beings (or If it is a mortal, then it is a human being)*

I concur with Zazkis and Chernoff (2008) that a conflicting example "does not have to be 'new,' but 'newly realized' or 'newly attended to'" (p. 196). For example, during the 2018-intervention, I used the statements *"If it is a person, then it is a mortal"* and *"If you are from Lima, then you are from Peru"* to create a conflict for Gessenia. Both statements were not "new" in Gessenia's background knowledge. Instead, they belonged to what Watson and Mason (2005) call her "example space". Zazkis and Chernoff (2008) suggest using such examples as "pivotal" examples[301]. My experience corresponded to this suggestion; Gessenia did not recognize any conflict when I presented a conflicting first example that was a mathematical statement, not in her "example space".

Ask the teachers whether St7 and St8 make the same claim, to identify the set of analysis for each statement, and whether St7 implies that St8 is true. Specially, draw attention to the set of analysis for each statement and request that the teachers provide examples of elements in the sets of analysis (see Section 3.1.5 above).

[300] See Chapter 5, Section I.1.1, for details.
[301] For an explanation on what a "pivotal example" and how it plays a role in the development of Gessenia's assumptions, see Chapter 5, Section I.1.1.2.

While statements St7 and St8 are a conflicting example with Assumption 5, they are a confirmatory example with Assumption 6. Introduce statements St9 and its converse St10 to avoid overgeneralizations that lead to Assumption 6.

St9: *If the unit-digit of a number is 0, then the number is divisible by 10*

St10: *If a number is divisible by 10, then its unit-digit is 0*

Statements St9 and St10 are a conflicting example since both are true statements. Focus the teachers' attention on the truth value of both statements. Ask the teachers whether the truth of St9 implies that St10 is also true. Draw attention to the set of analysis for each statement, request the teachers to provide examples of elements in the respective sets of analysis and ask them if it is possible to find elements that belong to one set and not to the other. Ask them whether it is possible to conclude something definite about the truth value of the converse of a true UAS.

Finally, return to an imaginary statement and its converse, such as St11 and its converse St12.

St11: *If it is a «poponupulus», then it is a «tataco»*

St12: *If it is a «tataco», then it is a «poponupulus»*

Ask them what could be inferred about the truth value of St12 if we assume that St11 is true. Draw the teachers' attention to elements that are *«tataco»* but not *«poponupulus»*. Ask them whether that case is possible or not and explain why. Instead of using new imaginary statements (St11 and St12), the initial confirmatory example could be used (i.e., St5 and St6).

7.1.4. All conditional statements are equivalent to universal statements

It is possible that some teachers make the generalization that all "if-then-statements" are equivalent to universal statements, as Andrea did during the 2018-intervention. Even though most "if-then-statements" can be expressed as universal statements, it is important that teachers encounter examples of "if-then-statements" that cannot be.

As outlined above, the discussion should begin with an example of a conditional statement that confirms the presumption (i.e., a conditional statement that is equivalent to a universal statement; e.g., St13). Attention should be drawn to this equivalence so that the teachers become aware of their existing assumptions. A second statement should be an example that contradicts the initial presumption; that is an example of a conditional statement that is not equivalent to a universal statement. For example:

St13: *If A is divisible by B, then B is divisible by A*

St14: *If A is divisible by B, then there are cases where B is divisible by A*

While St13 is equivalent to a universal statement, St14 is not. St14 is equivalent to an existential statement.

7.2. Leverage others' arguments to introduce valid and invalid forms of reasoning and submit them to criticism

Reactions, questions, assertions, hesitation, confidence, accuracies or inaccuracies, can reveal very much about the understanding process. In particular, both valid and invalid arguments that others have used to explain mathematical facts can support participants' understanding. Such arguments can be presented as new tasks, in the form of classroom episodes or arguments to be analyzed. For instance, for Discussion 3 I suggest including an activity that has been inspired by an interesting conjecture that a group of third graders came up with during my own teaching in 2015 (see Appendix A12). Because the mathematical content is the same, I used it in my 2017-intervention with teachers and they all wrongly accepted that the conjecture was true. This motivated me to include it again during the 2018-intervention and because of its effectiveness I suggest including it in Discussion 3. Similarly, in the 2018-intervention I included a discussion about the number of cases involved in a US, which was inspired by an inaccurate assumption that one of the teachers who participated in the 2017-intervention promoted during her teaching. I recommend including it in Discussion 4 since it was particularly relevant for Andrea' conceptualization of proof (see Chapter 5, Section I.3.1.1).

This view of teaching is in harmony with A. J. Stylianides and G. J. Stylianides' (2014) perspective that "at least part of prospective teachers' learning experiences should be contextualized in pedagogical situations" (p. 268). Further, they add that they use "a wide range of classroom scenarios based on actual classroom records: videos or written descriptions of classroom episodes in elementary classrooms, artifacts of elementary students' mathematical work, excerpts from elementary mathematics textbooks, etc. When actual classroom records were unavailable, we used fictional records that were nevertheless realistic." (p. 272). In that regard, the classroom context plays a crucial role as it grants access to a wide variety of forms of reasoning that participants use to reason mathematical issues.

Leveraging examples of valid and invalid arguments so that the teachers can discriminate which are mathematically valid justifications is an important approach. Teachers need to see representative examples of mathematically valid and invalid justifications and explain why they are valid or not from a mathematical point of view. My previous interventions made me aware of paradigmatic examples that were rich for proof-related discussions and introducing new forms of valid reasoning. In conducting an intervention, be attentive to interesting assertions and questions that can be used in future.

7.2.1. Use minimal and non-minimal valid justifications

Valid justifications can be distinguished from each other because of their sufficiency. For example, Discussion 2 has a focus on disproving USs, for which a minimal valid justification is one counterexample. During the corresponding discussion in the 2018-intervention, I made use of Lizbeth's sufficient or minimal justification to invite the teachers to discuss the validity of such an argument. This discussion was important in identifying some of the current Andrea and Gessenia's assumptions (see Chapter 5, Section I.2.1). It was also particularly relevant for Andrea who found her own reasons for why it made sense to consider it a sufficient justification (see Chapter 5, Section I.2.1.2). This shows that including a discussion where a counterexample is submitted to analysis is important to discussing sufficient justifications in the context of false USs. Likewise, this can motivate a discussion about the number of counterexamples that is sufficient to disprove a US. And a discussion about this issue can be used to introduce the case of non-minimal valid justifications.

Using a non-minimal justification that is also a mathematically valid justification to disprove a UAS is a good way to identify what aspects of a justification the teachers' attention is directed to. For instance, during the 2018-intervention Andrea was the only teacher who accepted the following non-minimal justification as valid to disprove the statement *"If a distribution is fair, whole and maximum, then the remainder is zero"*.

> *"The statement is not true, because if you have less objects than people it is indeed possible that we have one or more objects left."*

Andrea called this argument a *"verbalized counterexample"* (see Chapter 5, Section I.2.2.2). She recognized the argument as valid, which meant that her attention was placed on the characterization of the counterexamples that made the statement false. Given that this argument triggered an interesting discussion, I suggest using similar non-minimal justifications and submitting them to criticism in future interventions.

7.2.2. Include valid arguments with different forms of expression

During the 2018-intervention Gessenia initially rejected a counterexample as a valid justification to refute a UAS because of its form of expression. To her, it resembled a "calculation" and as we had previously used only narrative forms of expression to prove, this made her assume that a text form of expression was expected (see Chapter 5, Section I.2.1.1).

Including mathematical justifications with a wide variety of forms of expression supports a broadening of the teachers' perspectives about the different forms of expressions in which valid justifications can be presented. For example, a statement can be justified using different forms of expression, such as narrative, numerical (e.g., a counterexample, generic examples), algebraic, pictorial, etc. (see, e.g., Reid & Knipping, 2010).

7.2.3. Use also invalid justifications

It is important that teachers do not only analyze examples of valid justifications, but also invalid ones, and understand what aspects make them invalid.

During the 2018-intervention I used the following repetitive argument as an example of an invalid argument to disprove the UAS *"All divisions of natural numbers are exact divisions"*.

> *"It is false, because not all divisions of natural numbers are exact divisions"*

The argument was provided by one of the teachers who participated in the 2017-intervention. This was an important discussion during the 2018-intervention, in particular for Gessenia, who initially accepted a repetitive argument as valid (see Chapter 5, Section I.2.1.1). For a future intervention I suggest including a repetitive argument *before* discussing a sufficient justification. This is to identify whether teachers initially consider that repetitive arguments are valid justifications, without being influenced by a prior discussion on sufficient justifications. Discussing invalid justifications not only allows teachers to gain insights on what does not qualify as a valid mathematical justification, but also provides a basis for a discussion of the need to provide or include mathematically conclusive evidence.

7.2.4. Monitor the introduction of personal knowledge when evaluating arguments

It is crucial to develop awareness about the tricky issue of the introduction of personal knowledge in argumentation (see Morris, 2007). This is important in particular in the context of repetitive arguments (see Section 7.2.3). For example, during the 2018-intervention Gessenia assumed that a repetitive argument was sufficient to disprove UASs because she knew that there was a counterexample for the UAS under discussion and so she could use her personal knowledge to complete the argument. Gessenia introduced her own awareness of when the UAS was false and made her decision based on it. It is crucial to support the teachers' awareness that they should evaluate an argument as it was given, paying close attention not to add/remove information to/from it. This is also important when generic arguments (see Mason & Pimm, 1984; Reid & Vallejo-Vargas, 2018) are to be evaluated. In such contexts it might be especially difficult for a participant to leave aside her/his personal knowledge.

7.2.5. Be inspired by classroom discussions and create arguments to address specific issues

During the 2018-intervention Andrea rejected generic arguments at the beginning. She explained this by referring to the US *"For every* n *natural number (different from zero), the expression:* $1 + 1141n^2$ *does not produce a perfect square"* that we had previously analyzed, which had an extreme counterexample (see Chapter 5, Section I.3.1.1). She saw this case as showing that examples can *never* be used to prove universal statements about infinite sets. However, generic proofs include examples in which the main argument is based. They are very useful in primary school mathematics teaching as they grant access to deductive forms of reasoning (e.g., Stylianides et al., 2017). A limitation of the 2018-intervention was that I did not show the teachers a complete example of a generic argument that qualified as a proof. Instead, I pointed to general characteristics that made a generic argument a proof, which seemed to have partially persuaded Andrea. Nonetheless, teachers need to reflect and construct for themselves complete examples of generic arguments that qualify as valid proofs. They need to understand what makes those arguments valid proofs and become aware of the differences with mere confirming examples, which Andrea did not see at the beginning.

We can introduce valid and suitable modes of argumentation by presenting them as arguments made by others, which might be hard for the teachers to come up with for themselves. This was particularly useful with a justification for the maximal value of the remainder in a division of natural numbers, which served as a reference during the 2018-intervention (see Chapter 5, Section I.3.1.1). In other words, the analysis of others' arguments can support a reorientation of attention to the general reasons that are necessary for mathematically valid justifications.

I also suggest addressing an issue that arose during the 2018-intervention in relation to the disproving of existential statements. At one point Andrea assumed that existential statements were false because they had "counterexamples" (see Chapter 5, Section II.2). Including a similar argument can promote a discussion about what qualifies as a valid justification when disproving existential statements.

7.3. Use irrelevant and confirming examples to refine the characterization of counterexamples

Yopp (2017) highlights the importance of being able to construct a description of all possible counterexamples particularly when indirect reasoning is involved. In his article, Yopp focuses on contrapositive argumentation and notes that a crucial step for the Grade 8 participant in his study was to develop her awareness of those descriptions. Moreover, Yopp pointed out that his "work with Alison shows that students may benefit from explicit instruction about the distinction between constructing descriptions of mathematics objects and arguing whether or not these objects exist" (p. 165). Zaslavsky and Ron (1998) provided evidence that many of the students who were aware that counterexamples were sufficient to disprove a mathematical statement could not "distinguish between an example that satisfies the conditions of a counterexample and one that does not satisfy them" (p. 130).

Providing a description of all counterexamples to a universal statement is crucial when studying the truth value of the statement. Understanding a mathematical statement entails being aware of what it means for the statement to be true or false. In order to understand when a statement is true, it is necessary to understand what it means for it to be false, and vice versa (Epp, 2003). This is fundamental when identifying the conclusive evidence that dis/proves a statement. In addition, being aware of the characterization of counterexamples to a UAS facilitates a more conscious exploration of their non/existence. If they exist, then an individual can conclude that the statement is false by showing at least one counterexample; on the other hand, if they do not exist, the individual can conclude that the statement is true and s/he would need to argue about their non-existence (an episode related to the latter point is discussed in Chapter 5, Section I.3.3).

The incomplete input given to introduce what a counterexample is during the 2018-intervention led Gessenia to (incorrectly) assume that a counterexample should not satisfy at least one condition of the statement (see Chapter 5, Section I.2.2.3). In order to refine the initial given input (see Section 6.1.2 above), I suggest using challenging cases during Discussion 2, whose focus is on disproving USs. I introduced two types of challengers with that goal: irrelevant and confirming examples. Discussions should focus on whether those cases qualify as counterexamples or not and why. They are expected to be rejected as possible counterexamples and teachers should explain why those cases do not qualify as counterexamples, which supports finding a characterization for them (see Chapter 5, Section I.2.2).

Including confirming examples plays a crucial role because the initial input for *counterexample* ("A *counterexample for the statement S* is an example that contradicts *S*; an example that shows that *S* is false", see Section 6.1.2 above) explicitly states that it contradicts the statement, which should prompt the teachers to reject confirming examples as counterexamples. Likewise, including irrelevant examples should be rejected based on the teachers' understanding of the logical interpretation of universal statements (see Section 3.1 above). Because a US applies to all the elements in the set of analysis, the teachers should recognize that a counterexample should belong to the set of analysis, and as such irrelevant examples can be disregarded as counterexamples. Teachers should realize that because a counterexample contradicts a statement, and because it should satisfy the set-of-analysis condition, then a counterexample must contradict the second condition of the statement. That is, the role of both kinds of examples is to support the refinement of the characterization of *counterexamples* by discarding them as cases of counterexamples, based on the initial input for the meaning of *counterexample* and the teachers' understanding of the logical interpretation of a universal statement.

During the 2018-intervention, and particularly during the teachers' teaching, the teachers spontaneously used a similar approach to the one I used with them (see Chapter 5, Section I.2.2 or Chapter 6, Section III.1). Hence, I suggest using this approach in future interventions.

The suggestion I made in Section 4.1 (see above) could be also applied here to establish a general characterization of counterexamples. It can also facilitate the characterization of *hypothetical* counterexamples with the use of imaginary USs.

7.3.1. Use both types of irrelevant examples

Unlike confirming examples, there are two types of irrelevant examples. Neither of the two types satisfies the set-of-analysis condition of the statement, but they may or may not satisfy the consequent or conclusion condition of the statement. Including examples of both types will enrich the discussions. For example, discussing whether the irrelevant example 9 qualified as a counterexample for the conditional statement *"If a number is divisible by 6, then it is divisible by 3"* was important during the 2018-intervention. Observe that 9 does not satisfy the antecedent condition of the CS, but satisfies its conclusion condition and as such qualifies as a counterexample of the converse statement. It drew Lizbeth's attention to the *order in the statement* as an important factor to consider when producing a counterexample for it (see Chapter 5, Section I.2.2.1). This means that discussing this type of example can promote the recognition of the structure of statements, and also support tackling the problem of the converse (see Chapter 2, Section I.1).

7.3.2. Be aware and cautious of possible overgeneralizations

Above I mentioned Andrea's overgeneralization about using examples to prove. Here is a list of some other overgeneralizations I have observed that should be watched for and addressed if they occur.

A counterexample can disprove an ES

It is important to understand that counterexamples disprove universal statements, but there might be cases where the teachers overgeneralize that an existential statement is false because it has a "counterexample" (see Chapter 5, Section II.2). This potential issue can be tackled in Discussion 8 where the focus is on disproving ESs; however, it might arise already in Discussion 2, unless a special emphasis is given to the type of statements to which counterexamples apply; namely, universal statements.

A counterexample for a UNS has the same characteristics as those for a UAS

I suggest including a discussion on the falsity and disproving of UNSs as soon as the respective discussions for UASs is finished (see Discussion 2 in Table 31). To possibly avoid that the teachers overgeneralize that counterexamples for UASs and UNSs have exactly the same characteristics, I include five recommendations, some of which I have included before, but here they are applied to this particular case.

First, take into consideration that the teachers may already have initial assumptions about UNSs and their disproving. Do not forget to identify those initial assumptions when determining the organization of the upcoming discussions (see Section 4.2 above). During the 2018-intervention Andrea used a non-mathematical assumption for negative "all-statements" that influenced her decisions on their disproving (see Chapter 5, Section III.2.2).

Second, use a familiar form of UNS first. To avoid potential non-mathematical assumptions, I suggest using first a UNS of the form *"Every X is not Y"*, similar to what I suggested doing for the understanding of the logical interpretation of UNSs in Discussion 1. My goal is to alleviate the load that the form of a UNS might add to a first approach to determining its falsity and disproving it.

Third, vary the type of UNSs (see Section 4.1 above). Even if the tasks are similar, the goal is to increase the level of generality for the description of counterexamples. I suggest that at least three discussions be included: a first one involving a mathematical UNS; a second involving two imaginary statements (a UAS and a UNS); and a third discussion involving two abstract statements (a UAS and a UNS).

Fourth, draw attention to the understanding of the logical interpretation of UNSs (see Section 3.1 above). The three discussions I suggested in my previous recommendation should include a focus on showing understanding of the UNSs involved. In all three discussions there should be a request to *represent* the statements as an indirect way to exhibit this understanding. Discussion 1 already covers understanding the logical interpretation of UNSs; however, make sure that the teachers "activate" this understanding. The participants should use their insights when discussing a characterization of counterexamples to UNSs. During the 2018-intervention it was important to draw the teachers' attention to the understanding of the logical interpretation of a UNS and to providing confirming examples (i.e., examples that actually have the characteristics of counterexamples to a UAS with the same involved sets) (see Chapter 5, Section III.2.2.2). Include confirming and irrelevant examples and see whether the teachers can identify why those cases do not qualify as counterexamples for the UNSs in discussion.

Fifth, aim at an explicit contrast. As I suggested before, two discussions, where the main focus is on a UNS and the characterization of counterexamples to it, should include a UAS and a UNS. The type of statements might change, though they should involve the same sets. The aim is to make an explicit distinction between the descriptions of counterexamples to a UAS and a UNS. The second of these two discussions could also include a task with focus on whether a generic element is a counterexample for a UAS or for the respective UNS. For example, ask whether *n* is a counterexample for S9 or for S10, given that *n* is an element of *X* and *Y*. Ask the participants to explain their answers.

S9: *All X are Y*

S10: *All X are not Y*

During the 2018-intervention most of my attention on counterexamples was directed towards UASs. I did not focus on the description of counterexamples to UNSs. I suggest not making the same mistake in a future intervention.

7.4. Use indirect approaches to show the truth of universal statements and falsity of existential statements

Becoming aware of what it means for a statement to be true and what it means for it to be false is crucial. During the 2018-intervention I observed that this is particularly relevant in two cases; namely, when determining and proving the truth of a US and the falsity of an ES.

7.4.1. Draw attention to the impossibility finding counterexamples when proving USs

This approach is also known as eliminating the possibility of counterexamples (see Yopp, 2017). It consists in showing that a US, in particular of the form *"All X are Y"*, is true is the same as showing that the US has no counterexamples (i.e., showing that elements in *X* that are not in *Y* do not exist).

During the 2018-intervention Lizbeth began to develop a an assumption about *justification* that included this perspective (see Chapter 5, Section I.3.3). Lizbeth's emerging assumption made me aware that it may be plausible for primary school teachers to come to understand that proving that a US is true is also accessible through arguing why it is impossible for the US to be false.

I suggest including this form of reasoning in future interventions as it might emerge in Discussion 4, after discussing disproving USs in Discussion 2 and the status of statements in Discussion 3.

It is crucial that the teachers become aware of the need to show that counterexamples indeed do not exist. Draw the teachers' attention to this crucial point after directly asking them to consider whether the statement could be false. Emphasize that counterexamples may not always be reachable. For example, this is the case of the Conflicting Example 2 in Section 7.1.1 above. The message should be that sometimes counterexamples are not easy to find; however, this does not remove the possibility that they may exist. Hence, if they assert that there are no counterexamples for a US, they should show that there are not any with a valid justification. This form of reasoning can be introduced using examples of others' arguments (see Section 7.2). In Figure 68 I include an example of a proof eliminating the possibility of counterexamples for the statement *"Every prime number greater than 2 is not even"*.

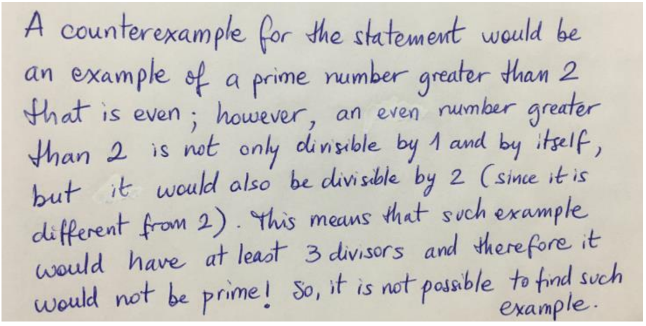

A counterexample for the statement would be an example of a prime number greater than 2 that is even; however, an even number greater than 2 is not only divisible by 1 and by itself, but it would also be divisible by 2 (since it is different from 2). This means that such example would have at least 3 divisors and therefore it would not be prime! So, it is not possible to find such example.

Figure 68. A proof as eliminating the possibility of counterexamples for the statement "Every prime number greater than 2 is not even".

This approach links a US and its negation. Furthermore, it shows the direct relationship between proving a US and disproving its negation. Showing that a US of the form *"All X are Y"* is true is the same as showing that its negation (i.e., *"Some X are not Y"*) is false; or, in other words, that it is false that there exists an *X* that is not *Y*. This approach may also be used as a way to discuss negations (see Section 9 below).

7.4.2. Draw attention to the impossibility finding confirming examples when disproving ESs

A similar form of reasoning was developed by Andrea when disproving existential statements during the 2018-intervention. She based her reasoning on her rationale that an ES was false given that it was impossible for the ES to be true (see Chapter 5, Section II.2). This involves that in order to disprove an existential statement of the form *"Some X are Y"*, it must be shown that it is impossible to find at least one *X* that is *Y*.

In this context, I suggest drawing the teachers' attention to the logical interpretation of existential statements. Ask them whether it is possible for the statement to be true and what kind of mathematical evidence they would need to show in that case. Guide them to reject that likelihood. Push the teachers to provide mathematical reasons that support that impossibility.

An example of a statement that the teachers disproved through their use of this approach was the ES *"Some divisions by 4 have a remainder that is equal to 7"*, for which they relied on the property that in divisions by 4, the maximal value of the remainder is 3.

These two suggestions (Sections 7.4.1 and 7.4.2) show the importance of the teachers becoming aware of what is involved in proving and disproving a statement S, when determining its truth value (Is it possible that the statement S is false? Is it possible that S is true?). The goal is that teachers spontaneously resort to this kind of dual analysis when the truth value of a statement is under discussion.

7.5. Pay attention to signs of emerging understandings to follow the course of the development of the assumptions

Evidence of the teachers' emerging understandings can be obtained by asking the teachers to express their insights from the discussions, both orally and in writing. They should share their insights explicitly and as clearly and accurately as possible. The teachers should be expressly prompted to explain the mathematical principles they identify. Some specific manifestations of emerging understandings are described in the following sub-sections.

7.5.1. Observe if the teachers cast doubt on their initial assumptions

Manifestations of doubt and hesitation about their initial assumptions are good signs of emerging understanding. For example, during the 2018-intervention, when discussing the statement *"If a number is even, then it is a «fast» number"*, Gessenia publicly questioned her initial assumption that the sets involved in a UCS have exactly the same elements. She asked *"Then, if a number is fast, it does not necessarily mean that it is even, or it does?... Does that mean that there are numbers that are not even and that are fast?"* (see Chapter 5, Section I.1.1.2)

On many occasions, uncertainty is evoked by conflicting statements, examples or situations (see Section 7.1 above). Dewey considered that "the origin of thinking is some perplexity, confusion or doubt" (Dewey, 1933, as cited in Zaslavsky, 2005, p. 299). Resolving their own uncertainties might be a motivation for the teachers' search for understanding. This is why paying careful attention to episodes where the teachers show uncertainty about their initial assumptions is relevant. They count as evidence of emerging understanding.

7.5.2. Observe if the teachers use more accurate language to refine their initial assumptions

Changes in the teachers' choice of words to express their assumptions more precisely is also a sign of new insights, more careful thinking and noticing. For example, during the 2018-intervention Andrea was more cautious about the way she modified one of her initial assumptions. Andrea put it as *if a UCS is true, its converse <u>may</u> be false,* instead of claiming that the converse *is* false (see Chapter 5, Section I.1.1.1). Likewise, Gessenia initially assumed that a UAS and its converse state the same thing. She refined her assumption to make the claim that *a UAS does not assert its converse, but it does not deny its converse either* (see Chapter 5, Section I.1.1.2). More accurate language use by the teachers is a positive sign of their emerging understandings.

7.5.3. Observe if the teachers produce examples with similar characteristics to conflicting ones

Teachers may notice a conflict and react to it by modifying their initial assumption; even better, however, is when they move beyond that and attempt to further explain their insights by producing examples that have the same characteristics as those from a conflicting case. To do this, teachers must have identified what made those cases conflicting in the first place. For example, during the 2018-intervention, Gessenia produced an example of a UAS with a false converse to show that she understood that a UAS does not necessarily assert the truth of its converse (see Chapter 5, Section I.1.1.2). Similarly, Andrea came up with two examples of "some-statements" that were universally true, which conflicted with her initial assumption that true "some-statements" could not be universally true because of her initial personal meaning for "some" (see Chapter 5, Section II.1).

In both of these cases, the teachers spontaneously provided examples; however, if participants do not do this, explicitly ask for them. For the teachers to notice the characteristics of the conflicting examples and create their own examples requires a certain degree of abstraction. This already suggests an implicit awareness of and an agreement with the contradicting information that is part of a cognitive conflict used to provoke change in the teachers' initial assumptions.

7.5.4. Observe if the content influences emerging assumptions

New assumptions should be tested, for instance, by shifting the content in which the statements are framed. If the teachers change their new assumptions based on the change of content, it is likely that their assumptions are content dependent. The aim is that they form context-independent assumptions.

In Section 3.1.1 (above) I pointed to this as an important factor to consider when drawing the teachers' attention to the logical interpretation of a statement, but it is also relevant here where attention is on the teachers' emerging assumptions.

7.5.5. Observe if emerging understandings change when using tasks where different mathematical activities are the focus

Use tasks involving different activities (e.g., evaluating a given argument, providing an argument). This may support the identification of initial assumptions as well as helping to detect possible repeated use of those assumptions.

For example, during the 2018-intervention Gessenia used her initial assumption that *repetitive arguments were sufficient arguments to disprove a UAS* in two different contexts. First, a task requested her to evaluate a given (repetitive) argument for a false UAS; and second, a task expected her to provide a justification to refute a UAS (she gave a repetitive argument). The second task was presented after we had discussed why repetitive arguments are not valid arguments (see Chapter 5, Section I.2.1.1). In this particular case, the first task made me aware of Gessenia's initial assumption, while the second task let me observe whether she paid attention to this kind of invalid argument in order to not to use them again. This shows how including tasks involving different mathematical activities can allow testing for the recurrence of initial assumptions.

7.5.6. Pay attention to the teachers' explanations

The teachers' explanations can reveal their current understandings, but also their confusions, doubts or lack of understanding.

Do not assume that the other teachers will use an accepted explanation from one teacher to reason about related issues. It is possible that they do not attend the same factors or that because they are starting with other assumptions the explanation has a different meaning for them. For example, during the intervention Andrea and Lizbeth explained that a counterexample should satisfy the first condition, but contradict the second condition of the US under discussion. Despite that, Gessenia still seemed confused about this and continued to assume that a counterexample should contradict the statement as a whole, without making precise what exactly should be contradicted in the statement (see Chapter 5, Section 2.2.3).

Observing the teachers' teaching is another way to notice which explanations from the intervention they understood and include in their own teaching, and which they continue to struggle with. If possible, work as a participant observer, occasionally providing inputs in the class, especially when the teachers lack confidence in leading the proof-related discussions (see Chapter 4, Section II.2.2). This can also be a way to identify the teachers' current understandings. Observe the aspects or elements of statements that the teachers highlight and pay attention to. This allows you to identify the factors that play an important role in the teachers' personal insights. Their teaching may be a crucial source of information about what assumptions they use in each moment. Paying close attention to the teachers' explanations not only gives access to their understandings, but also to their limitations (for examples, see Chapter 6, Section II).

7.5.7. Read the teachers' personal insights diaries

If you have requested the teachers to keep a "diary" to write down their insights and their respective explanations, you should read those diaries and provide personal feedback.

While expressing their emerging understandings orally to the whole class can be a good way to share insights and agree on the clearest and most accurate way to express them, keeping an individual written record and receiving personal feedback, is also important. My experiences with teachers in general, and with the teachers in the 2018-intervention in particular, showed me that teachers normally do not feel confident expressing their thinking orally when proof-related discussions are involved. Keeping a record of the teachers' writing and providing them feedback can be a way to support the teachers' improvement in expressing mathematical principles and their explanations. It can also be decisive to their engagement in oral discussions.

If the teachers struggle with writing or expressing their ideas clearly, this will probably be revealed in their diaries and you will be able to know who needs support and what kinds of support they might need.

II. Topic specific design principles

In this section I focus on principles related to "no-statements" and negations. Unlike the principles in Section I, these principles are based on problems I experienced during the 2018-intervention and my reflections on them. In contrast to the principles in Section I that were tested in at least one of the two design cycles, these principles are provisional and offered here as suggestions of possible ways to address the problems I encountered.

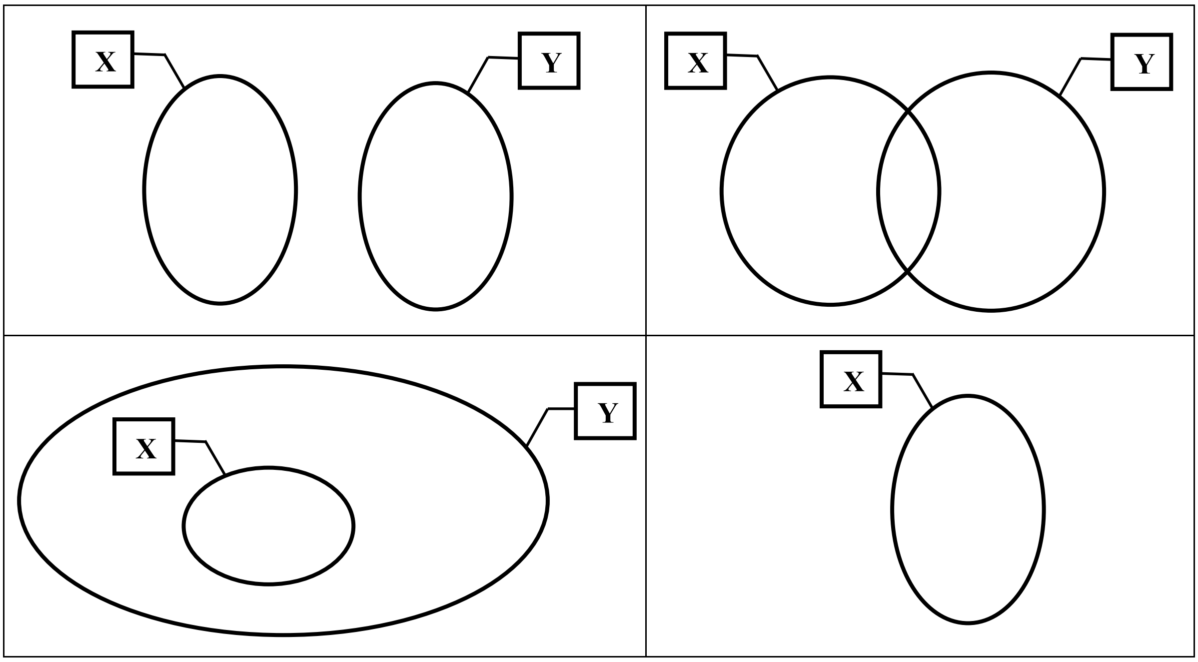
8. Principles for the block about "no-statements"

In Section 2.1 I suggested beginning the discussions with focus on universal statements (see Table 31). Even though "no-statements" are universal statements, I have decided to discuss them separately, based on my experience with the teachers during the 2018-intervention. Specifically, Andrea did not relate "no-statements" to "all-statements". The way she interpreted "no-statements" was far from its mathematical interpretation (see Chapter 5, Section III.1.1). Based on this experience I believe it is important to focus on "no-statements" in a separate block in future.

8.1. Begin with a simple negation: "Not X"

Ask the teachers to shade the area that represents *"not X"* in the following diagrams (see Table 32). It is important to find out what the teachers initially assume about this specific issue before engaging them in discussions about "no-statements". During the 2018-intervention Andrea interpreted the statement *"No X is Y"* as referring to the elements that were *not X*, which she presumed were the elements in *Y*, according to the first diagram below (the diagram in the first row and first column in Table 32).

Table 32. Diagrams where "not X" is expected to be shaded.



8.2. Promote the link between "no-statements" and "all-statements"

Getting to know the teachers' initial assumptions is a step I suggest is crucial for each discussion (see Section 4.2). In particular, it is relevant to find out how teachers interpret "no-statements", which includes whether the teachers grasp that "no-statements" refer to all the elements in their set of analysis.

Understanding the following equivalence should be a goal:

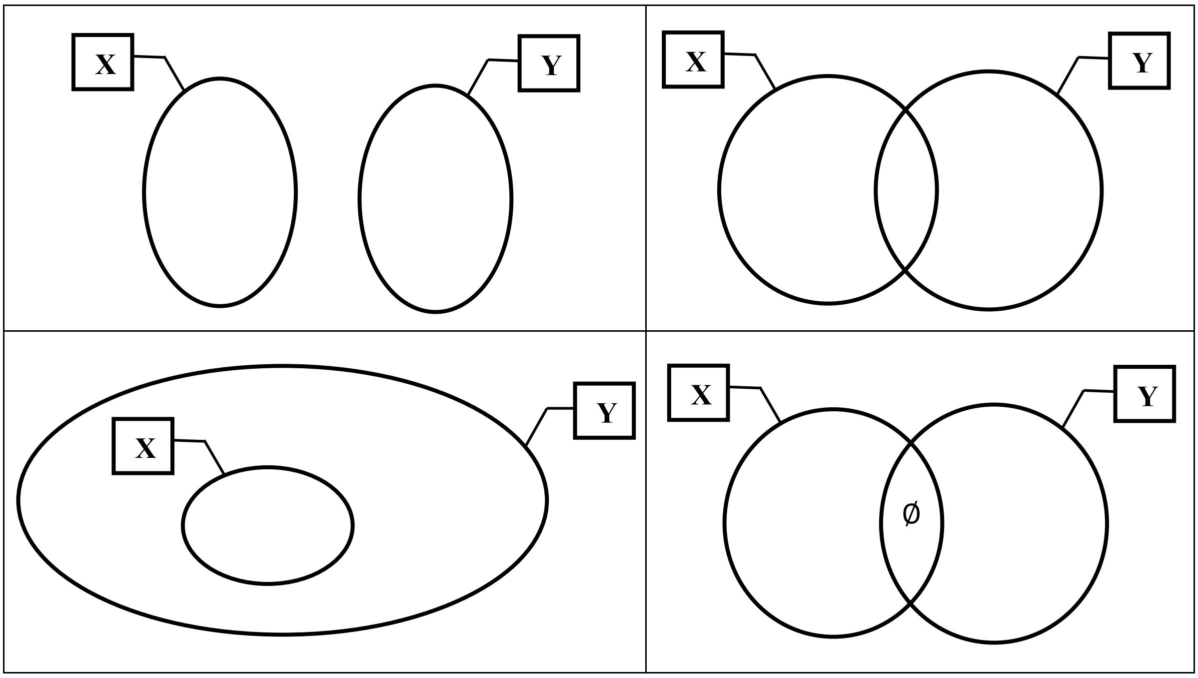
"No X is Y" ≡ "All X are not Y"

One approach to reaching understanding of this relation is asking the teachers to choose which of the following four diagrams in Table 33 represent the claim made in *"No X is Y"* and to explain why[302]. However, this may not be a sufficient approach as the teachers can identify a suitable diagram, but still keep their initial assumption that "no X…" refers to everything that is not X, as Andrea did during the 2018-intervention (for details, see Chapter 5, Section III.1.1). The teachers should be clear about the logical interpretation of "no-statements", and notably that "No X…" refers to all elements in X, before reasoning further about "no-statements".

Once this relationship is established, then reasoning about "no-statements" can be based in an appropriate understanding of "all-statements". Hence, it is very important to promote that linkage.

[302] Observe that the difference between the two diagrams in the second column is that while the diagram in the second row explicitly states that the intersection is empty, the diagram in the first row does not. This should invite the teachers to a discussion of what to assume when nothing is explicitly stated in a diagram.

Table 33



9. Principles for the block about Negations

> *"At a conceptual level, when complex logical statements are used to define a concept, it may be said that in order to understand what something is, it is essential to understand what it is not."*
>
> (Dubinsky, Elterman & Gong, 1988, p. 46)

Statements are connected. We cannot isolate a topic like universal statements from existential statements, given that they are strongly related through negation and therefore disproving. As a statement is proved to be false, its negation is indirectly proved to be true and, vice versa, when a statement is proved to be true, its negation is proved to be false. Primary school teachers' knowledge about this link (or lack of it) may be unconscious and revealed, for instance, when discussing single-quantified statements. During the 2018-intervention this was particularly clear in the case of Andrea, who assumed that the *semantic substitution* (Dawkins, 2017) for "not all" and "there does not exist" were "some" and "none", respectively. Andrea initially used these semantic substitutions to find equivalent statements for implicit negations of both "all-statements" and "there-exist-statements", respectively (see Chapter 5, Sections I.4 and II.3.2). The fact that this primary school teacher used initial assumptions about the connection suggests that interventions with focus on proof-related issues should include discussions about negations. Nonetheless, I include this topic as the last block in the design for a future intervention (see Table 31) because my expectation is that teachers understand non-negated statements first and then use this understanding as a basis for their understanding of negations.

In the following sub-sections I consider some concrete suggestions for the topic of negation and how I would organize it in a future intervention, according to my experience with the 2018-intervention.

9.1. Develop an initial sense of negative categories

There is a common procedure the teachers used when dealing with negative categories during the 2018-intervention and which should be considered in future interventions. They converted negative categories into positive categories. For example, during the first part of the intervention Andrea changed "not even" to "odd". These are *immediate contraries* (Horn, 2001) that are exactly equivalent. However, the teachers also converted the category of numbers "not greater than zero" into the category of numbers "smaller than zero", rather than "smaller or equal to zero". The teachers were not aware that the relation "not greater than zero" allowed an intermediate case (zero) and thus "not greater than zero" and "smaller than zero" were *mediate contraries* (Horn, 2001, p. 7). Other examples of mediate contraries are "white" and "black" since grey might be a color in between, or "good" and "bad". This is one of the important features of negative categories to discuss.

Some authors have highlighted the importance of reasoning with negative categories when engaging in proof-related discussions in mathematics.

> Mathematical logic depends heavily upon the universal partition into examples and non-examples… Mathematicians correspond negative properties to the complement of the set designated by the property negated, which I call the *negation/complement relation*. (Dawkins, 2017, p. 502)

I did not pay enough attention to this issue during the 2018-intervention, and I suggest it be a focus in any future intervention, especially the negation/complement relation and the understanding of the logical interpretation of "no-statements" (see Section 8 above).

9.2. Use implicit negations first

Postpone discussions about *explicit* negations until after discussions that promote an understanding of *implicit* negations. By *implicit negation* I mean those negations that are already included in the statement itself (e.g., *"**Not** all puppies are mischievous"* or *"**It is not the case that** some even numbers are palindrome numbers"*). A *simple implicit negation* is formed by adding "no" or "not" to the statement (e.g., *"**Not** all puppies are mischievous"*) (for details on explicit and implicit negations, see Chapter 5, Section I.4).

Focus first on supporting an understanding of what it means for a statement to be false as well as finding different equivalent ways to express falsity. In this regard, Sellers (2018) has suggested that "the word 'negation' may be more appropriate to use after students have constructed formal rules for negation based on their justifications for why statements are false" (p. 254). Nonetheless, we need to be aware that there might be cases where teachers could introduce the term "negation" or a related one themselves. For instance, Andrea explicitly emphasized that "not" in a statement of the form *"Not all X are Y"* played the role of a negator (see Chapter 5, Section I.4). In cases like this I believe that it is important to discuss negations, but the first step is to find out whether the teachers are aware of what exactly the negator "not" negates in the statement, and to identify possible non-mathematical assumptions. For example, Andrea used her initial assumption that "not" affected only the quantifier "all" and because of her semantic substitution, "not all" could be switched with "some".

Because some teachers might already hold initial assumptions about implicit negations (see Chapter 5, Sections I.4 and II.3), I suggest including implicit negations next and finding out their initial assumptions. For example, ask them to find an equivalent statement for a statement of the form *“Not all X are Y”* or *“There does not exist X that are not Y”*. Are they aware that those are negations?

One way to introduce statements that the teachers might not come up with on their own is including multiple-choice tasks, where you ask the teachers to choose those options that make the same claim as the main (implicit negation) statement. Figure 69 is an example of such a prompt for such a discussion.

Discussion 9A

Consider statement S1:

S1: Not all «boguis» are «nodees»

Choose all the options that include a statement that makes exactly the same claim or means the same as statement S1.

a) All «boguis» are not «nodees»
b) Some «boguis» are «nodees»
c) There are «boguis» that are «nodees»
d) No «boguis» are «nodees»
e) There are no «boguis» that are «nodees»
f) There exist «boguis» that are not «nodees»
g) Some «boguis» are not «nodees»
h) If it is a «boguis», then it is not «nodees»

Figure 69. A discussion prompt for finding out the teachers’ initial assumptions about implicit negations of “all-statements”

9.3. Support an explicit establishment of the linkage between implicit negations and falsity

Linking the implicit negation of a statement and its falsity is related to the *negation of the meaning* that Dubinsky, Elterman and Gong (1988) discuss. It is mainly explaining the negation of single-quantified statements by resorting to their falsity.

Do not introduce negation as a rule. Rules make the teachers rely merely on a mechanical procedure to find negations and do not challenge the teachers to make sense of why they work the way negations do. Instead, draw the teachers’ attention to the meaning of the implicit negation of a statement by relating it to what it means for the statement to be false. It might be useful to analyze the statements you used before when determining the teachers’ initial assumptions (e.g., Statement S1 in Figure 69) or to use a harmonious statement (e.g., *“Not all people are mean”*). Link the meaning of the implicit negation *“Not all «boguis» are «nodees»”* with what it means for the statement *“All «boguis» are «nodees»”* to be false. You could also rely on the following equivalences:

Not all «boguis» are «nodees» = It is not the case that all «boguis» are «nodees» = It is not true that all «boguis» are «nodees» = It is false that all «boguis» are «nodees»

In Table 34 (below) , I include a sequence of tasks that could guide a discussion about this linkage between implicit negations and falsity. I use abstract universal statements to point to a general statement whose form can be changed.

Task 1 directly addresses a common non-mathematical assumption that many individuals are prone to use when dealing with implicit negations of UASs (e.g., see Pasztor & Alacaci, 2005) where the contradictory is confounded with the contrary (in terms of the *Traditional Square of Oppositions*, see Horn, 2001). In Figure 69 I also considered the identification of this possible initial assumption (see Option d). Task 2 aims at checking whether the teachers hold a mathematically-aligned assumption for the falsity of USs. Task 3 already puts forward the link between implicit negation and falsity and whether the teachers are aware of that linkage. Task 4 has a focus on determining whether the teachers are aware of the mathematical evidence that shows the falsity of USs. Task 5 draws attention to the teachers' awareness of the truth value for the negation of a statement based on the given truth value of the statement. Task 6 aims to draw attention to the mathematical evidence needed to prove an implicit negation (Option b). This is requested by first retrieving the teachers' understanding of disproving USs (Option a), as a way that the teachers use such awareness to prove the equivalent implicit negation, based on the equivalence relation between the implicit negation and falsity (see Task 3). Task 7 has a focus on equivalent forms to express an implicit negation.

Table 34. Sequence of generic tasks suggested for Negation

T#	Task	Focus of task
1	Is *"Not all X are Y"* the same as *"No X is Y"*?	A common non-mathematical assumption about negations of "all-statements": *"Not all X are Y"* is the same as *"No X is Y"*.
2	What does it mean that the statement *"All X are Y"* is false?	Current understanding of falsity of UASs.
3	Are the two following statements (St1 and St2) related? St1: *Not all X are Y* St2: *It is false that all X are Y* How? What about St3? Does it relate to St1 and St2? St3: *It is not the case that all X are Y*	Linkage between the simple implicit negation of a UAS and its falsity. Alternative forms of implicit negation.
4	What do you need to show to disprove *"All X are Y"*?	Conclusive evidence to disprove a US.
5	If *"All X are Y"* is false, a) what can you say about the truth value of the statement *"Not all X are Y"*? Is it true or false? Why? b) What about the truth value of *"It is false that all X are Y"*? c) What now if *"All X are Y"* is true?	Link between the truth value of a US and the truth value of its implicit negation.
6	a) What do you need to show to prove that *"It is false that all X are Y"*? b) What do you need to show to prove that *"Not all X are Y"*?	Conclusive evidence to prove falsity. Use of the link between negation and falsity to find out what is involved in proving an implicit negation.
7	St1: *Not all X are Y* Pick the statements that have the same interpretation as St1. a) *Some X are Y* b) *It is not true that all X are Y* c) *There is X that is not Y* d) *All X are not Y* e) *Some X are not Y* f) *It is false that all X are Y* g) *It is not the case that some X are Y* h) *No X are Y*	Equivalent statements for the implicit negation of a US.

Take into consideration that at this point the teachers are supposed to be already aware of the falsity and disproving of single-quantified statements (I suggest negations be encountered in Discussion 9, much later than discussions about falsity, see Table 31). In other words, my expectation is that the teachers' current understanding of disproving a statement will play a crucial role when discussing the negation of those statements. For the example in Figure 69, I assume that the teachers can manage to make connections between the implicit negation of the UAS and the existential statement that is its negation; namely, *"Some «boguis» are not «nodees»"*. Being aware of the characteristics of the possible counterexamples that may refute a UAS (see Discussion 2) can support the teachers' identification of the negation of UASs.

There are different levels of awareness that the teachers should learn about. Are the teachers aware that a statement such as *"Not all X are Y"* is indeed a negation? Are they aware of what is concretely negated in such statement? Are they aware of what it means the negation of a statement? Are the teachers aware about the relation between negation and falsity? Are they aware of equivalent forms to express negation?

9.4. Postpone the explicit request to negate statements as much as possible

Once implicit negations are understood, and their linkage with falsity is established, then the path towards explicit negations has better foundations. Task 8 in Table 35 (below) explicitly addresses the issue of making explicit an implicit negation. In particular, it has a focus on the role that “not” plays in the statement *“Not all X are Y”*, which may trigger a discussion about negations. Task 9 supports the teachers’ awareness that *“Not all X are Y”* is the negation of the statement *“All X are Y”*. Task 10 aims at establishing the link between falsity and explicit negations. Task 11 brings out the equivalence between the implicit negation of a US and an ES as a way to establish a relation between an explicit negation of the US and an ES. Task 12 aims at establishing the link between the explicit negation of a US and an ES.

Table 35. Continuation of Table 34

T#	Task	Focus of task
8	a) What is the role that “not” plays in the statement *“Not all X are Y”*? b) What does it negate precisely?	The role that “not” plays in the implicit negation of a US.
9	Is *“Not all X are Y”* the negation of any statement? Which one?	Establishing an explicit negation: *“Not all X are Y”* is the negation of *“All X are Y”*.
10	Complete the following sentences in such a way that it becomes a true sentence. *“It is false that all X are Y” = “Not ___X_____Y” =* The negation of the statement *“ __________ ”*	The link between falsity and explicit negations.
11	Choose the suitable word(s) so that the following become true sentences. a) *“Not all X are Y” = “(Some/All/No) ________X (is/is not/are/are not) ________ Y”* b) The negation of *“All X are Y” = “(Some/All/No) _______ X (is/is not/are/are not) ________ Y”*	Implicit and explicit negations of a US. The link between USs and ESs through negation.
12	Is the negation of the statement *“All X are Y”* a universal or an existential statement? Which one?	The link between the explicit negation of a US and an ES.

To sum up, my suggestion is to, first, make sure that the teachers understand what is involved in disproving a statement; then, address negations by focusing on the particular case of implicit negations, without necessarily pointing out that they are negations, unless that is suggested by one of the teachers. Only after a strong connection between implicit negations and falsity of statements has been established, would I recommend that explicit negations are discussed. Drawing the teachers’ attention to the role that “not” plays in an implicit negation might be a path to begin to discuss explicit negations. Discuss explicit negations by making a linkage with implicit negations and the understandings built for those cases. Transfer those understandings to the case of explicit negations.

In the same line, Sellers (2018) suggested that the request to explicitly negate statements should be used after the students have elaborated a rule for negation that is based on their justifications for why a statement is false. So, more important than avoiding the use of a word (“negation” in this case) is that the teachers first become aware of the relation between implicit negation and falsity, and make of that a relation that can be easily brought out and applied in contexts where it is useful. For example, it can be used when determining equivalent statements to express negation and/or falsity in different forms in order to understand a statement, find out its truth value, justify it, etc.

9.5. Present the task of negating a statement in different equivalent forms

Presenting the negation of a statement in different forms (e.g., in its implicit and explicit forms) might also support connections that should be established when discussing the negation of statements in mathematics. This supports bringing the teachers' conscious attention to negation and the many forms used to express it.

Ask the teachers to choose which of the following tasks make the same request.

1) Find an equivalent statement for *"There does not exist X that is Y"*
2) Find an equivalent statement for *"It is not the case that some X are Y"*
3) Find the negation of the statement *"There does exist X that is Y"*
4) Find the negation of the statement *"Some X are Y"*

Alternatively, you can ask the teachers to provide as many equivalent forms as possible to negate a specific statement.

9.6. Establish the link between proving and disproving of implicit negations

To find out whether the teachers easily understand the proving and disproving of implicit negations, you can include some tasks to determine the teachers' initial assumptions about this issue before any linkage between implicit negations and falsity is established. After that you can contrast their initial assumptions with a more conscious understanding based on that connection.

Showing that an implicit negation is true or false may be straightforward in some cases. For example, during the 2018-intervention the teachers seemed to not struggle with three particular cases: disproving simple forms of implicit negations of EASs (e.g., disproving *"There does not exist X that is Y"*), proving simple forms of implicit negations of UASs (e.g., proving *"Not all X are Y"*), and disproving affirmative "no-statements" (e.g., disproving *"No X is Y"*). Nevertheless, cases of negative USs, ESs and "no-statements" were more challenging (see Chapter 5, Sections I.4, II.3 and III.1.3).

A future intervention should relate the following three processes with regard to, for example, a UAS of the form *"All X are Y"*: Proving that *"Not all X are Y"*, proving that *"It is false that all X are Y"*, and disproving *"All X are Y"*. The teachers should come to understand that those three processes are equivalent; that is,

Proving that *"Not all X are Y"* = Proving that *"It is false that all X are Y"*
= Disproving *"All X are Y"*

Likewise, the teachers should be able to understand the following relation:

Disproving that *"Not all X are Y"* = Disproving that *"It is false that all X are Y"*
= Proving *"All X are Y"*

Becoming aware of the relation between implicit negations and falsity is needed to develop a conscious understanding of their proving and disproving. It is possible that in some cases (like the ones I pointed out above) the teachers' performances are spontaneous; however, it is still crucial that they become aware of why those requests are equivalent.

9.7. Draw attention to simplifications of proving and disproving of explicit negations

The case of explicit negations and their truth value should not be difficult if the teachers manage to understand first that once a statement is proved to be true or false, its negation is implied to have the opposite truth value. The teachers can just focus on the main statement, determine its truth value, and then provide the opposite truth value for the negation.

You can draw attention to the request to prove that the negation of the US is false (e.g., Prove that the negation of *“All X are Y”* is false) is the same as proving that the US should be proved to be true. Another approach could be finding the negation of the US (i.e., *“Some X are not Y”*) and disproving the resulting statement.

Prove that the negation of *“All X are Y”* is false =

$$= \text{Prove that } \underbrace{\text{the negation of } \underbrace{\textit{“All X are Y”}}_{\text{True}}}_{\text{False}} \text{ is false}$$

= Prove *“All X are Y”*

Likewise, prove the negation of *“There exist X that is Y”* =

$$= \text{Prove } \underbrace{\text{the negation of } \underbrace{\textit{“There exist X that is Y”}}_{\text{False}}}_{\text{True}}$$

= Disprove *“There exist X that is Y”*

= Prove *“There does not exist X that is Y”*

To determine the truth value of the explicit negation of a statement, a similar form of reasoning can be used. The teachers’ attention can be first drawn to the main statement for which truth value should be changed in order to find out the truth value of the negation. For instance, if the truth value of the negation of *“Some X are Y”* is being discussed, attention can be directed towards the truth value of the main statement *“Some X are Y”* and once it is determined, the truth value of its negation can be concluded to be the opposite.

9.8. Guide the teachers to negate statements as a whole and to not divide them

It is important that the teachers understand that cutting the statement into parts and interpreting the parts in order to make sense of the whole statement might be misleading. During the 2018-intervention, Andrea did this (see Chapter 5, Sections I.4 and III.1.1).

Be aware that teachers might make conversions such as *“Not all X are Y”* is the same as *“Some X are Y”* because “not all” is the same as “some” (see Chapter 5, Section I.4). To avoid this, clarify that “not” actually negates the whole statement *“All X are Y”* and as such it negates the whole structure. Ask the teachers to explain what it means for the statement *“All X are Y”* not to be true. If possible, have them use representations such as Venn or Euler diagrams. Ask the teachers to explain what it means that the representation of the US *“All X are Y”* is false; draw their attention to the falsity that all the elements of set *X* belong to set *Y*.

9.9. Be cautious of double negations

It is important to detect whether teachers interpret double negations as single negations, especially if double negations are used for emphasis, as in Spanish. For example: do they initially interpret the statement *"There does not exist X that is not Y"* as *"There does not exist X that is Y"* or *"No X is Y"* as the teachers did in my 2018-intervention? (see Chapter 5, Section II.3). Consider ways to tackle the possible assumption that double negations can be interpreted as single ones in mathematics. Resorting to the logical interpretation of the statement and the fact that the "main" negation affects the whole statement (in the example it is the negation of the statement *"There exist X that is not Y"*) might have potential. For example, ask the teachers to consider both statements *"There exist X that is not Y"* and *"There does not exist X that is not Y"*, and explain the differences they notice between them. Ask them to provide equivalent statements for both so that you can verify whether they interpret them mathematically. You can have the teachers to represent each statement by using Venn or Euler diagrams and explore their interpretations.

III. Conclusion

In this chapter I provided some design principles, in the form of suggestions for an intervention to engage teachers with mathematical forms of reasoning. They are an outcome of my reflections on the 2018-intervention. Both positive and negative effects of the intervention played an important role in my formulation of these recommendations. I believe that they would be useful for planning interventions with a similar focus on proof and proof-based teaching.

Chapter 8: Suggestions for Future Research and Reflections

In my research I addressed the question of how to introduce proof to primary school teachers who had no previous experience with it. Nevertheless, more needs to be said and done about this issue. In this chapter I include some questions that might shed some light on directions for future research. I structure those questions according to three main themes: design, proof understanding and transfer.

I. Design of interventions

I did the intervention twice. From the second intervention, I drew principles that I included in Chapter 7, which I would like to explore in a new round of design-based research with a new group of teachers.

In my interventions I included examples of dis/proof that supported the teachers' refinements of their conceptualization of proof. In particular, I restricted the statements we discussed to single-quantified statements and so the cases applied to such a frame. In that context, the teachers' conceptualization of proof was gradually and continuously refined. This is not very far from the way many of us learned how to prove. Mathematicians seem to similarly "update" their standards to accept or not certain "proofs" (e.g., the "proof" of the four-color theorem, a computer assisted proof). Krantz (2011) claims that "it is essential that the mathematical community have a formal recognition of the changing and developing nature of mathematical proof" (p. 226). Nonetheless, is this an accessible and effective way to present proof to other teachers who are also novices in that field? What is important for teachers to understand about proof in order to engage themselves and engage others in proof-related activities?

The first part of the 2018-intervention was focused on the mathematical content. My goal was that in the second part the teachers could mostly (if not only) focus on the development of mathematical forms of reasoning. In that respect, using a mathematical content that was framed in a PfBT theory was relevant (see Chapter 1, Chapter 4 and Reid & Vallejo-Vargas, 2019). It controlled the set of definitions and axioms to be used as the basis for the teachers' arguments, which in a way allowed everyone in our community to better understand each other. So, what if I had not included the first part of the intervention (mathematical content) and I had directly focused on proof-related discussions? What if it was not content framed in a PfBT theory? Would the effects of the intervention be similar to the ones I observed? To what extent should that mathematical content be developed?

A. J. Stylianides and Ball (2008) have highlighted that the mathematical knowledge that teachers need to design their own teaching sessions with a focus on proof-related issues has to be investigated. The design of the 2018-intervention addressed this need. It included two parts: the mathematical content and proof-related discussions. The teachers were expected to teach the same mathematical content and engage their students in similar proof-related discussions. The sequence of their teaching sessions resembled the organization of the intervention and was submitted to them for approval before they implemented them in their classrooms. The teachers could suggest changes for their teaching sessions. The design of the 2018-intervention mitigated the planning load rather than leaving this responsibility completely on the teachers' shoulders. At the beginning this might be an important step to support the teachers' confidence in teaching a new and non-trivial topic like "proof"; however, identifying when the appropriate moment for the teachers to be in complete charge of designing their teaching sessions with a focus on proof might be tricky. Is the approach I took going to make the teachers negatively dependent on an initial given sequence in the long term? Can an invitation to suggest one or two proof-related tasks and a critical discussion about them be a way to support the teachers' future independence? These questions require further research.

During the 2018-intervention I made most things explicit (e.g., whether one counterexample was sufficient to disprove a universal statement; whether a counterexample was sufficient to disprove an existential statement; in what cases confirming examples were sufficient to prove a universal statement). However, not all the teachers' conceptualizations of proof were expressly emphasized. I tried to explicitly highlight the evolution of the teachers' assumptions and I recommend to have open discussions to keep a common record of all the participants' agreed understandings, besides the teachers' personal insights (e.g., see Chapter 7, Section I.7.5.7). Nonetheless, it is also true that the teachers did not always gain insight of the same forms of reasoning involved in our discussions. For instance, sometimes one teacher did not find evident what other teacher might have (e.g., Andrea's acceptance of non-minimal proofs to disprove USs). So, would making things even more explicit allow a uniformization of the paths taken by the teachers' understandings? Is such uniformization a desirable goal? If so, what can make those paths match in such a way that the understandings become common to all members of that community?

The 2018-intervention included several opportunities where provoking cognitive conflicts played an important role in changing the teachers' initial assumptions. However, there were some specific aspects that supported its usage and effect. In the context of the 2018-intervention, it was crucial to tackle the most fundamental assumptions in order to have higher chances for effective cognitive conflicts. In this respect, Gal (2019) identified that the problematic learning situations reported in her article show that none of the teachers who participated in the study tried to find out the origins of their students' struggles, which did not allow them to make a more informed decision to handle them. The case of Gessenia and the converse problem during the 2018-intervention (see Chapter 5, Section I.1.1.2) showed me that it is possible that the most fundamental initial assumption is accessed much later during the intervention. In my view, that accounted for the fact that some of her assumptions did not change even after using cognitive conflicts. Cognitive conflicts that are focused on initial assumptions that are based on other, more fundamental, assumptions may not trigger the expected change. Real change might come only when the most fundamental assumption is addressed. Is a cognitive conflict approach more effective when the most fundamental assumption is tackled? Assuming that the assumptions a teacher holds about a certain topic (e.g., valid inferences from a true universal statement) are linked to each other, how can one determine which one is the most fundamental? What if those assumptions are not seen as related to each other by the participants? Should those links be made explicit at some point during an intervention? All these questions need to be investigated.

II. Proof understanding

In Chapter 6, Section III, I included the teachers' assumptions that changed as revealed during their teaching. In the three cases the teachers resorted to the logical interpretation of statements to support their pupils' insights and it became a common practice. Is the understanding of the logical interpretation of statements a fundamental understanding required to address further proof-related understandings? Mostly, students are requested to construct a proof for a given statement, without first exploring whether their interpretation of the statement is compatible with the mathematical interpretation. Nevertheless, this can be a factor that influences students' performance when dis/proving a mathematical statement and needs to be investigated. G. J. Stylianides and A. J. Stylianides (2020) point out other factors that should not be overlooked when doing research on this field. One of them is to consider the individuals' own perceptions of the arguments they produce and whether they qualify as proof or not when the focus is on investigating the individuals' justification schemes. However, even in that case, I believe that it is important to explore what teachers initially understand by "proof" (or "justification") and some other related terms; otherwise, why should we be sure that they will show something that matches our expectations? This gets more difficult to achieve if those individuals have no previous experience with dis/proving and their main source of assumptions originates from the daily life context, as I showed in my research (see e.g., Chapter 5, Sections I.1, I.4 and II.1).

One feature of the 2018-intervention is that it allowed the teachers to have opportunities to express their arguments both in writing and orally. A. J. Stylianides (2019) pointed to both *modes of argument representation* (A. J. Stylianides, 2007) to be considered when studying individuals' construction of arguments that need to meet the standard of proof. In particular, this was important to my intervention as a way to perform a more complete and accurate analysis of the development of the teachers' proof-related assumptions. Furthermore, the justifications they produced and evaluated mainly relied on a text form of expression, both oral and written, and on calculations they could handle. Only a few of them were algebraic, which not all the teachers immediately followed. Does that mean that primary school teachers cannot cope with algebraic justifications? Probably not, but this suggests that, as in any process, teachers need to develop their ability to justify algebraically progressively. In any case, given that primary school teachers may not be familiar with different forms of valid (deductive) modes of argumentation, a first approach to address generality can be built up by engaging teachers in using generic proofs. How feasible is that primary school teachers actively engage in such mode of argumentation to consciously target generality? Is it possible that as they understand that confirming examples may not prove an infinite universal statement they become skeptical about generic examples when included in generic proofs? Recall that Andrea experienced this skepticism during the 2018-intervention (for details, see Chapter 5, Section I.3.1.1). Is Andrea's skepticism about generic examples a rare case or might it be a common reaction when learning about mathematically valid methods to prove?

During the 2018-intervention the teachers spontaneously began to develop indirect forms of reasoning. For example, Lizbeth began to develop an assumption about proving a US as the impossibility to find counterexamples to it (for details, see Chapter 5, Section I.3.3) and Andrea developed an assumption about disproving a "some-statement" as showing the impossibility for it to be true (for details, see Chapter 5, Section II.2). Further research needs to be done in terms of identifying what it takes to develop a need to provide a general argument to refute the existence of counterexamples when proving a US (e.g., in Lizbeth's case) or the existence of confirming examples when disproving an ES (e.g., in Andrea's case). Likewise, further research needs to be done to investigate whether they are viable forms of reasoning that other teachers would spontaneously develop or adopt when dis/proving SQ-statements.

Some of the teachers' initial assumptions were not in harmony with mathematical ones. In some cases, those assumptions were uncommon. For example, Andrea's initial assumption that a true UAS implied a false converse is an uncommon assumption (see

Chapter 6, Section I.1)[303]. Is this assumption also used by other teachers and maybe it is not that unusual? Furthermore, it is strange that despite the non-mathematical interpretation of some statements, related assumptions were mathematically accurate. For instance, Andrea's initial logical interpretation of "no-statements" was not aligned with its counterpart in mathematics; however, several related assumptions were (e.g., their disproving, their representation, their negation, see Chapter 6, Section I.4). Does that mean that a mathematical interpretation of "no-statements" is unnecessary to further mathematically reason about them? In the same vein, Andrea's initial assumptions about negations involved her use of semantic substitutions to find equivalent statements for simple implicit negations such as *"There does not exist X that is Y"* and *"Not all X are Y"*. The question is again whether these assumptions are only specific to Andrea, or in any case, how spread are they in reality? Do other teachers hold the same initial assumptions? It is possible that we have been ignoring them even though they also deserve attention. Research needs to be developed in order to find out what is the case. These cases enlighten our field about forms of reasoning used by primary school teachers that may affect their proof-related assumptions and their teaching when engaging their students in proof-related discussions.

In terms of the development of the teachers' assumptions, there might be cases in which the teachers' initial non-mathematical assumptions change to mathematical assumptions; however, the nature of those changes might not be mathematical (see Chapter 6, Section I.4). For example, Andrea's initial approach to negate "all-statements" changed from *"Not all X are Y"* is the same as *"Some X are Y"* to *"Not all X are Y"* is the same as *"Some X are not Y"*; however, she based her new assumption on a shortcut that was the result of her combination of a rule I gave to negate universal statements and her initial semantic substitution approach, which did not reveal any form of understanding (for details, see Chapter 5, Section I.4). Investigating the nature of changes in the teachers' assumptions should be addressed in future research that has a focus on proof-related issues.

In their study, Mejía-Ramos et al. (2012) investigated a model for assessing proof comprehension. Within the aspects that their model assesses they include: meaning of terms and statements, logical status of statements and proof framework, justification of claims, summarizing via high-level ideas, identifying the modular structure, transferring the general ideas or methods to another context, illustrating with examples. The authors associated the three first aspects with a *local* comprehension of a proof, whereas the other four were gathered to assess a *holistic* comprehension of a proof. This paper came to my attention after I developed my research; however, I can see many similarities between Mejía-Ramos et al.'s (2012) model and my work with primary school teachers. Even though the authors focused on assessing proof comprehension when *reading* proofs and I focused on the aspects that teachers should pay attention to when engaged in proving, in both cases the main aim is understanding proof. On a macro scale, in Mejía-Ramos et al.'s (2012) terms, I mainly focused on a local comprehension of proof. Further research needs to be done to find out whether a local comprehension of proofs in fact suffices to have a holistic comprehension of proof.

Language is an important issue that has shown to be relevant in proof-related research (see e.g., Buchbinder & McCrone, 2020) and particularly in my research. In my work I have shown that the daily-life language influences teachers' proof-related assumptions about single-quantified statements in mathematics. For example, it influenced the

[303] For details, see Chapter 5, Section I.1.1.1.

teachers' understanding of the logical interpretation of statements. Hence, an intervention that intends to focus on similar goals to mine needs to consider the interpretation of SQ-statements in the daily-life context language, which includes the meaning of quantifiers. In addition, this should consider the language that is spoken in the country as well as the challenges it adds to the logical interpretation of SQ-statements. As I have shown with my research, some teachers might hold mathematically-sound proof-related assumptions despite their non-mathematical interpretations of the statements. Andrea's interpretation of "no-statements" was a clear example of this, which did not affect other proof-related assumptions such as disproving "no-statements" (for details, see Chapter 5, Section III.1). What if the logical interpretation of statements is not addressed before engaging the teachers in proof-related discussions about those statements? Is it possible that teachers understand dis/proving without becoming aware of their initial interpretation of statements and an appropriate attention to it? The main aim of focusing on the logical interpretation of statements was that the teachers have a consistent set of proof-related assumptions. So far, I have not seen that previous research had a concern on this aspect: *consistency* in the set of students' proof-related assumptions. Is paying careful attention to the logical interpretation of statements a (the only) way to achieve that goal?

Antonini and Mariotti (2008) used the expression "meta-theorem" for a notion that includes a "meta-proof" for a mode of argumentation; however, their focus is on sophisticated ways to prove a mode of argumentation (e.g., by using a formal logic approach), which according to Yopp (2017) might not be accessible to average school students or might not be effective, as Antonini's (2004) study shows. In the 2018-intervention, the teachers justified several assumptions about modes of argumentation by basing their explanations on the logical interpretation of the statements. For example, Andrea explained why one counterexample was sufficient mathematical evidence to infer that a universal statement was false by resorting to the logical interpretation of the universal statement (for details, see Chapter 5, Section I.2.1). Do these kinds of explanations qualify as "meta-proofs", in Antonini and Mariotti's sense, for modes of argumentation? Do they support better the teachers' modes of argumentation than the sophisticated ways to prove them that Antonini and Mariotti consider as meta-proofs? Are they more effective? These questions need to be investigated.

III. Transfer

The 2018-intervention was developed to promote the understanding of proof-related forms of reasoning that govern mathematical reasoning and dis/proving. My focus was on arithmetic, and specifically on *division and divisibility*, as this is a topic that is taught at primary school level. Dawkins and Karunakaran (2016) have pointed to the mathematical content as an influential factor of students' proof-related behavior. Whether Andrea, Gessenia and Lizbeth will still use their insights when reasoning a different mathematical content (e.g., algebra or geometry content) is uncertain. My hypothesis is that given that their understanding of the logical interpretation of mathematical statements was foundational, this understanding would still be present when they reason single-quantified statements for other mathematical contents; however, it might be limited by a possible lack of familiarity with the mathematical content itself. This needs to be further investigated.

It is important to investigate whether changing the form of the universal statement would change the teachers' assumptions about the characterization of counterexamples for it. For example, investigate if the same counterexamples are proposed for a universal conditional statement presented as an *"X if Y"* structure (e.g., "a number is divisible by 3 if it is divisible by 6") rather than the more common "if-then" structure (e.g., "If a number is divisible by 6, then it is divisible by 3"). Is it possible that the teachers transfer their understanding of "all-statements" to other universal statements with different form by making connections among USs with different forms of expression? In the same vein, research needs to be done in order to explore whether refining the concept "counterexample" gradually for the case of universal *affirmative* statements results in a conscious characterization of counterexamples to universal *negative* statements. Andrea seems to have done that for the case of negative "for-every-statements" (for details, see Chapter 5, Section III.2.1). Is it possible that the teachers adapt and transfer their understanding of universal affirmative statements to universal negative statements to enlarge their sets of mathematical assumptions?

IV. Reflections

All in all, I have achieved the goals I had set up for my work. I designed an intervention with a focus on supporting the development of primary school teachers' proof-related understandings and, notably, their understanding of the nature of proving (Chapters 4 and 5). I investigated the effects the intervention had on the teachers' abilities to engage in proof-related activities and how it is reflected while they teach in their classrooms in a proof-based way (Chapters 5 and 6). To do that I paid close attention to the assumptions that changed and assumptions that did not change but the reasons that supported them did change. I identified the features of the intervention that may explain changes in the teachers' proof-related understandings (see Chapter 6).

I made a contribution towards addressing what basic principles primary school teachers need to understand about proof and proof-related issues (Chapter 7). I have underscored that instead of cutting up the analysis of the development of the teachers' proof-related assumptions to see specific things, it would be important to engage in research that aims at analyzing the teachers' assumptions from a global perspective. Certainly, there is more research to be done in this area. In chapter 7 I made some suggestions for developing interventions for teachers including aspects that need to be attended and in the previous sections I gave some directions for future related research that still needs to be done.

I focused on how primary school teachers' assumptions change as they engage in an intervention focused on Proof-Based Teaching and understanding the nature of proving and what they need to understand to engage themselves and their pupils in dis/proving. The answer might be different to secondary school teachers, or it might be different if the attention is on a mathematical topic different from division and divisibility. I feel this work provides a solid basis for future research on teachers' proof competencies and developing them through interventions for pre- and in-service teachers.

References

Abrusci, M., Pasquali, F., & Retoré, C. (2016). Quantification in ordinary language and proof theory. *Philosophia Scientae*, *20*(1), 185–205. https://doi.org/10.4000/philosophiascientiae.1163

Alacaci, C., & Pasztor, A. (2005). People's Negation Behavior: The Power of Content. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (Vol. 27, p. 2440). Hillsdale: Erlbaum.

Albertí, M. (2011). *La creatividad en matemáticas: Cómo funciona una mente maravillosa*. Barcelona RBA.

Anderson, J. R. (2015). *Cognitive Psychology and its Implications* (8th edition). Worth Publishers, New York, NY.

Antonini, S. (2001). Negation in mathematics: obstacles emerging from an exploratory study. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 49–56). The Netherlands, Utrecht.

Antonini, S. (2004). A statement, the contrapositive and the inverse: Intuition and argumentation. In M. Johnsen Høines, & A. Berit Fuglestad (Eds.), *Proceedings of the 28th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 47–54). Norway, Bergen.

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: What is specific to this way of proving? *ZDM*, *40*(3), 401–412. https://doi.org/10.1007/s11858-008-0091-2

Baggini, J., & Fosl, P. S. (2010). *The Philosopher's Toolkit: A compendium of Philosophical Concepts and Methods* (2nd Ed.). Chichester, UK. Wiley-Blackwell.

Bakker, A., & van Eerde, D. V. (2015). An introduction to Design-Based Research with an example from statistics education. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods* (pp. 429–466). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9181-6_16

Balacheff, N. (1986). Cognitive versus Situational Analysis of Problem-Solving Behaviors. *For the Learning of Mathematics*, *6*(3), 10–12. http://www.jstor.org/stable/40247818

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, Teachers and Children* (pp. 216–235). London: Hodder and Stoughton.

Balacheff, N. (1991). Treatment of Refutations: Aspects of the Complexity of a Constructivist Approach to Mathematics Learning. In E. von Glasersfeld (Ed.), *Radical Constructivism in Mathematics Education* (pp. 89–110). Springer, Dordrecht. https://doi.org/10.1007/0-306-47201-5_5

Balacheff, N. (2002). The researcher epistemology: A deadlock for educational research on proof. In F.L. Lin (Ed.), *Proceeding of the 2002 International Conference on Mathematics: Understanding proving and proving to understand* (pp. 23–44). Taipei, Taiwan: NSC and NTNU.

Balacheff, N. (2008). The role of the researcher's epistemology in Mathematics education: An essay on the case of proof. *ZDM*, *40*(3), 501–512. https://doi.org/10.1007/s11858-008-0103-2

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 27–44). Reston, VA: NCTM.

Ball, D. L., Hoyles, C., Jahnke, H. N., & Movshovitz-Hadar, N. (2002). The Teaching of Proof. In L. I. Tatsien (Ed.), *Proceedings of the International Congress of Mathematicians* (Vol. III, pp. 907–920). Beijing, China: Higher Education.

Bardelle, C. (2013). The Influence of Verbal Language in a Mathematical Context: A Study on Negation. In A. M. Lindmeier, & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 65–72). Kiel, Germany: PME.

Barkai, R., Tsamir, P., Tirosh, D., & Dreyfus, T. (2002). Proving or Refuting Arithmetic Claims: The Case of Elementary School Teachers. In A. D. Cockburn, & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 57–64). Norwich, England: PME.

Barnard, T. (1995). The Impact of 'Meaning' on Students' Ability to Negate Statements. In L. Meira, & D. Carraher (Eds.), *Proceedings of the 19th International Conference of the International Group for the Psychology of Mathematics Education.* (Vol. 2, pp. 3–10). Recife, Brazil: PME.

Boom, J. (2009). Piaget on equilibration. In U. Müller, J. Carpendale, & L. Smith (Eds.), *The Cambridge companion to Piaget* (pp. 132–149). Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL9780521898584.006

Brousseau, G. (2002). *Theory of didactical situations in mathematics: Didactique des Mathématiques, 1970–1990*. Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47211-2

Bucci, W. (1978). The interpretation of universal affirmative propositions. *Cognition*, 6, 55–77. https://doi.org/10.1016/0010-0277(78)90009-4

Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *Journal of Mathematical Behavior, 58*, 100779. https://doi.org/10.1016/j.jmathb.2020.100779

Buchbinder, O., & McCrone, S. (2022). Guiding principles for teaching mathematics via reasoning and proving. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti. (Eds.). *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education* (CERME12). Free University of Bozen-Bolzano, Italy, and ERME. https://hal.archives-ouvertes.fr/hal-03746878v2

Buchbinder, O., & Zaslavsky, O. (2007). How to decide? Students' ways of determining the validity of mathematical statements. In D. Pitta-Pantazi, & G. Philippou (Eds.), *Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education* (pp. 561–570). Larnaca, Cyprus: University of Cyprus.

Buchbinder, O., & Zaslavsky, O. (2009). A framework for understanding the status of examples in establishing the validity of mathematical statements. In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 225–232). Thessaloniki, Greece.

Buchbinder, O., & Zaslavsky, O. (2011). Is this a coincidence? The role of examples in fostering a need for proof. *ZDM*, *43*(2), 269–281. https://doi.org/10.1007/s11858-011-0324-7

Buchbinder, O., Zodik, I., Ron, G., & Cook, A. L. (2017). What can you infer from this example? Applications of online, rich-media tasks for enhancing pre-service teachers' knowledge of the roles of examples in proving. In A. Leung, & A. Baccaglini-Frank (Eds.), *Digital Technologies in Designing Mathematics Education Tasks* (pp. 215–235). Cham: Springer. https://doi.org/10.1007/978-3-319-43423-0_17

Cabassut, R., Conner, A., İşçimen, F., Furinghetti, F., Jahnke, H., & Morselli, F. (2012). Conceptions of Proof –In Research and Teaching. In G. Hanna, & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: the 19th ICMI Study* (pp. 169–190). New York: Springer.

Campbell, R. (2009). Constructive Processes: Abstraction, Generalization, and Dialectics. In U. Müller, J. Carpendale, & L. Smith (Eds.), *The Cambridge Companion to Piaget* (pp. 150–170). Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL9780521898584.007

Carroll, L. (2009). *Alice's Adventures in Wonderland and Through the Looking-Glass*. Oxford University Press.

Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, *15*(1), 1–40. https://doi.org/10.1207/s1532690xci1501_1

Chazan, D. (1990). Quasi-empirical views of mathematics and mathematics teaching. *Interchange, 21*(1), 14–23. https://doi.org/10.1007/BF01809606

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, *24*(4), 359–387. https://doi.org/10.1007/BF01273371

Chinn, C. A., & Brewer, W. F. (1993). The role of Anomalous Data in Knowledge Acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*(1), 1–49. https://doi.org/10.3102/00346543063001001

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher*, *32*(1), 9–13. https://doi.org/10.3102/0013189x032001009

Copi, I. (1973). *Introducción a la lógica* (14va ed.). Buenos Aires: Eudeba.

Copi, I. M., Cohen, C., & McMahon, K. (2014). *Introduction to Logic* (14th Ed.). Pearson.

Cosmides, L. (1989). The Logic of Social Exchange: Has natural selection shaped how humans reason? Studies with the Wason Selection Task. *Cognition*, *31*(3), 187–276. https://doi.org/10.1016/0010-0277(89)90023-1

Czocher, J. A., & Weber, K. (2020). Proof as a cluster category. *Journal for Research in Mathematics Education*, *51*(1), 50–74. https://doi.org/10.5951/jresematheduc.2019.0007

Dahlberg, R. P., & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics, 33*, 283–299. https://doi.org/10.1023/A:1002999415887

Davis, P. J., & Hersh, R. (1998). *The Mathematical Experience*. Houghton Mifflin Harcourt.

Dawkins, P. C. (2017). On the importance of set-based meanings for categories and connectives in Mathematical Logic. *International Journal of Research in Undergraduate Mathematics Education*, 3(3), 496–522. https://doi.org/10.1007/s40753-017-0055-4

Dawkins, P. C., & Cook, J. P. (2017). Guiding reinvention of conventional tools of mathematical logic: Students' reasoning about mathematical disjunctions. *Educational Studies in Mathematics*, *94*(3), 241–256. https://doi.org/10.1007/s10649-016-9722-7

Dawkins, P. C., & Karunakaran, S. S. (2016). Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content. *The Journal of Mathematical Behavior*, *44*, 65–75. http://dx.doi.org/10.1016/j.jmathb.2016.10.003

Dawkins, P. C., & Roh, K.H. (2016). Promoting Metalinguistic and Metamathematical Reasoning in Proof-Oriented Mathematics Courses: A method and a framework. *International Journal of Research in Undergraduate Mathematics Education, 2*(2), 197-222. https://doi.org/10.1007/s40753-016-0027-0

Dawkins, P. C. (2019). Students' pronominal sense of reference in mathematics. *For the Learning of Mathematics*, *39*(1), 13–18.

Dawkins, P. C., & Roh, K. H. (2022, February 2-6). The role of unitizing predicates in the construction of logic [Conference presentation]. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti. (Eds.). *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education* (CERME12). Free University of Bozen-Bolzano, Italy, and ERME. ⟨hal-03746875v1⟩

Deloustal-Jorrand, V. (2007). Relationship between beginner teachers in mathematics and the mathematical concept of implication. In D. Pitta-Pantazi, & G. Philippou (Eds.), *Proceedings of the fifth congress of the European Society for Research in Mathematics Education*. Larnaca: CERME 5.

Dreyfus, T. (2000). Some views on proofs by teachers and mathematicians. In A. Gagatsis (Ed.), *Proceedings of the 2nd Mediterranean Conference on Mathematics Education* (Vol. I, pp. 11–25). Nicosia: The University of Cyprus.

Dubinsky, E., Elterman, F., & Gong, C. (1988). The Student's Construction of Quantification. *For the Learning of Mathematics*, *8*(2), 44–51. http://www.jstor.org/stable/40247924

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics*, *53*(1), 5–34. https://doi.org/10.1023/A:1024661004375

Durand-Guerrier, V. (2008). Truth versus validity in mathematical proof. *ZDM*, *40*(3), 373–384. https://doi.org/10.1007/s11858-008-0098-8

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. S., & Tanguay, D. (2012). Examining the role of logic in teaching proof. In G. Hanna, & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: The ICMI Study* (pp. 369–389). New York: Springer. https://doi.org/10.1007/978-94-007-2129-6_16

Durand-Guerrier, V., Kazima, M., Libbrecht, P., Ngansop, J. N., Salekhova, L., Tuktamyshov, N., & Winsløw, C. (2016). Challenges and opportunities for second language learners in undergraduate Mathematics. In R. Barwell et al. (Eds.), *Mathematics Education and Language Diversity: The 21st ICMI Study* (pp. 85–101). Cham: Springer. https://doi.org/10.1007/978-3-319-14511-2_5

Durand-Guerrier, V. (2020). The logical analysis of statements. A tool for dealing with ambiguities in multilingual context. In J. Ingram, K. Erath, F. Rønning, & A. K. Schüler-Meyer (Eds.), *Proceedings of the Seventh ERME Topic Conference on Language in the Mathematics Classroom* (pp. 31–38). Montpellier, France: ERME. ⟨hal-02970549⟩

Duval, R. (2007). Cognitive functioning and the understanding of mathematical processes of proof. In P. Boero (Ed.), *Theorems in school: From history, epistemology and cognition to classroom practice* (pp. 137–161). Rotterdam, the Netherlands: Sense Publishers. https://doi.org/10.1163/9789087901691_009

Ellis, A. B., Lockwood, E., Dogan, M. F., Williams, C. C., & Knuth, E. (2013). Choosing and Using Examples: How Example Activity Can Support Proof Insight. In A. M. Lindmeier, & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 265–273). Kiel, Germany: PME.

Epp, S. S. (1998). A unified framework for proof and disproof. *The Mathematics Teacher, 91*(8), 708–713. http://www.jstor.org/stable/27970751

Epp, S. S. (1999). The language of quantification in mathematics instruction. In L. V. Stiff, & F. R. Curcio (Eds.), *Developing mathematical reasoning in grades K-12 (1999 Yearbook)* (pp. 188–197). Reston, VA: National Council of Teachers of Mathematics.

Epp, S. S. (2003). The Role of Logic in Teaching Proof. *The American Mathematical Monthly, 110*(10), 886–899. https://doi.org/10.1080/00029890.2003.11920029

Epp, S. S. (2009a). Proof issues with existential quantification. In F.L. Lin, F.J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of ICMI study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 1, pp. 154–159). Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.

Epp, S. (2009b). The use of logic in teaching proof. In B. Hopkins (Ed.), *Resources for Teaching Discrete Mathematics: Classroom Projects, History Modules, and Articles (MAA Notes)* (pp. 313–322). Mathematical Association of America. doi:10.5948/UPO9780883859742.038

Epp, S. S. (2020). *Discrete Mathematics with Applications* (5th Ed.). Cengage Learning, Boston.

Evans, J. S. B. T. (1972). Interpretation and Matching Bias in a Reasoning Task. *Quarterly Journal of Experimental Psychology*, *24*(2), 193–199. https://doi.org/10.1080/00335557243000067

Evans, J. S. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human Reasoning. The Psychology of Deduction*. Lawrence Erlbaum Associates, Inc.

Ferrari, P. L. (2004). Mathematical language and advanced mathematics thinking. In M. Haines & A. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 383–390). Bergen, Norway: Bergen University College.

Fischbein, E., & Kedem, I. (1982). Proof and certitude in the development of mathematical thinking. In A. Vermandel (Ed.), *Proceedings of the Sixth International Conference of the Psychology of Mathematics Education* (pp. 128–131). Antwerp, Belgium: Universitaire Instelling Antwerpen.

Freeman, N. H., & Stedmon, J. A. (1986). How children deal with natural language quantification. In I. Kurcz, G. W. Shugar, & J. H. Danks (Eds.), *Knowledge and Language* (Vol. 39, pp. 21–48). Amsterdam: Elsevier. https://doi.org/10.1016/S0166-4115(09)60131-X

Gal, H. (2019). When the use of cognitive conflict is ineffective—problematic learning situations in geometry. *Educational Studies in Mathematics*, *102*(2), 239–256. https://doi.org/10.1007/s10649-019-09904-8

Giannakoulias, E., Mastorides, E., Potari, D., & Zachariades, T. (2010). Studying teachers' mathematical argumentation in the context of refuting students' invalid claims. *The Journal of Mathematical Behavior*, *29*(3), 160–168. https://doi.org/10.1016/j.jmathb.2010.07.001

Goff, W., & Getenet, S. (2017). Design-based research in doctoral studies: Adding a new dimension to doctoral research. *International Journal of Doctoral Studies*, *12*, 107–121. https://doi.org/10.28945/3761

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics: Speech Acts* (pp. 41–58). New York: Academic Press.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational studies in Mathematics*, *44*(1), 127–150. https://doi.org/10.1023/A:1012781005718

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Baltimore: University Park Press.

Hanna, G., & Jahnke, H. N. (1996). Proof and Proving. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education. Kluwer International Handbooks of Education* (Vol. 4). Springer, Dordrecht. https://doi.org/10.1007/978-94-009-1465-0_24

Hanna, G. (2000). Proof, Explanation and Exploration: An Overview. *Educational studies in mathematics*, *44*(1), 5–23. https://doi.org/10.1023/A:1012737223465

Hanna, G. (2018). Reflections on Proof as Explanation. In A. Stylianides, & G. Harel (Eds.), *Advances in Mathematics Education Research on Proof and Proving. ICME-13 Monographs* (pp. 3–18). Springer, Cham. https://doi.org/10.1007/978-3-319-70996-3_1

Harel, G. (2001). The Development of Mathematical Induction as a Proof Scheme: A Model for DNR-Based Instruction. In S. Campbell & R. Zaskis (Eds.), *Learning and Teaching Number Theory*. Journal of Mathematical Behavior (pp. 185–212). New Jersey: Ablex Publishing Corporation.

Harel, G., & Sowder, L. (1998). Students' Proof Schemes: Results from Exploratory Studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education III* (Vol. 7, pp. 234-283). Providence, RI: AMS, CBMS: Issues in Mathematics Education. https://doi.org/10.1090/cbmath/007/07

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805–842).

Hawkins, J., Pea, R. D., Glick, J., & Scribner, S. (1984). "Merds that laugh don't like mushrooms": Evidence for deductive reasoning by preschoolers. *Developmental Psychology*, *20*(4), 584–594. https://doi.org/10.1037/0012-1649.20.4.584

Healy, L., & Hoyles, C. (2000). A Study of Proof Conceptions in Algebra. *Journal for Research in Mathematics Education*, *31*(4), 396–428. https://doi.org/10.2307/749651

Hempel, A. & Buchbinder, O. (2022, February 24-26). Undergraduate students' use of everyday language to make sense of indirect proof [Poster presentation]. *24th Conference on Research in Undergraduate Mathematics Education* (RUME), Boston, United States. NO URL Yet.

Horn, L. R. (2001). *A Natural History of Negation*. Standford: Center for the Study of Language and Information (CSLI) Publications.

Hoyles, C., & Küchemann, D. (2002). Students' understandings of logical implication. *Educational Studies in Mathematics*, *51*(3), 193–223. https://doi.org/10.1023/A:1023629608614

Iannone, P., Inglis, M., Mejía-Ramos, J. P., Simpson, A., & Weber, K. (2011). Does generating examples aid proof production?. *Educational studies in Mathematics*, *77*(1), 1–14. https://doi.org/10.1007/s10649-011-9299-0

Jahnke, H. N., & Wambach, R. (2013). Understanding what a proof is: a classroom-based approach. *ZDM*, *45*(3), 469–482. https://doi.org/10.1007/s11858-013-0502-x

Jones, K. (1997). Student-teachers' conceptions of mathematical proof. *Mathematics Education Review, 9*, 21–32.

Kempen, L. (2018). How Do Pre-service Teachers Rate the Conviction, Verification and Explanatory Power of Different Kinds of Proofs? In A. J. Stylianides, & G. Harel (Eds.), *Advances in Mathematics Education Research on Proof and Proving* (ICME-13 Monographs, pp. 225–237). Springer, Cham. https://doi.org/10.1007/978-3-319-70996-3_16

Knuth, E. J. (2002a). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education*, *33*(5), 379–405. https://doi.org/10.2307/4149959

Knuth, E. J. (2002b). Teachers' conceptions of proof in the context of secondary school mathematics. *Journal of Mathematics Teacher Education*, *5*(1), 61–88. https://doi.org/10.1023/A:1013838713648

Knuth, E. J., Choppin, J. M., & Bieda, K. N. (2011). Middle school students' production of mathematical justifications. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *Teaching and Learning Proof across the grades: A K-16 perspective* (pp. 153–170). New York, NY: Routledge.

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Weigel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 1693-1670). Athens, GA: Clearinghouse for Science, Mathematics, and Environmental Education.

Ko, Y. Y., & Knuth, E. J. (2009a). Undergraduate mathematics majors' writing performance producing proofs and counterexamples about continuous functions. *The Journal of Mathematical Behavior*, *28*(1), 68–77. https://doi.org/10.1016/j.jmathb.2009.04.005

Ko, Y.Y., & Knuth, E. (2009b). Problems manifested in prospective secondary mathematics teachers' proofs and counterexamples in differentiation. In F.L. Lin, F.J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of ICMI study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 1, pp. 262–267). Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.

Ko, Y. Y. (2010). Mathematics teachers' conceptions of proof: Implications for educational research. *International Journal of Science and Mathematics Education, 8*(6), 1109–1129. https://doi.org/10.1007/s10763-010-9235-2

Krantz, S. G. (2011). *The proof is in the pudding: The changing nature of mathematical proof*. New York: Springer.

Lakatos, I. (1976). *Proofs and Refutations*. Cambridge: Cambridge University Press.

Lee, K., & Smith, J. P., III (2009). Cognitive and linguistic challenges in understanding proving. In F. L. Lin, F. J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of ICMI study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 21–26). Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.

Lee, K. (2016). Students' proof schemes for mathematical proving and disproving of propositions. *The Journal of Mathematical Behavior*, *41*, 26–44. https://doi.org/10.1016/j.jmathb.2015.11.005

Lew, K., & Zazkis, D. (2019). Undergraduate mathematics students' at-home exploration of a prove-or-disprove task. *The Journal of Mathematical Behavior*, *54*, 100674. https://doi.org/10.1016/j.jmathb.2018.09.003

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, *11*(4-5), 357–380. https://doi.org/10.1016/S0959-4752(00)00037-2

Lin, F. L., Lee, Y. S. & Wu, J. Y. (2003). Students' understanding of proof by contradiction. In N.A. Pateman, B.J. Doughterty & J.T. Zilliox, (Eds.), *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 443–450). Honolulu, Hawaii: PME.

Lin, P. J., & Tsai, W. H. (2012, July 8–15). Future teachers' proof of universal and existential statements. Paper presented at the *12th International Congress on Mathematical Education* (ICME-12, TSG 23). COEX, Seoul, Korea.

Lockwood, E., Ellis, A. B., & Lynch, A. G. (2016). Mathematicians' example-related activity when exploring and proving conjectures. *International Journal of Research in Undergraduate Mathematics Education*, *2*(2), 165–196. https://doi.org/10.1007/s40753-016-0025-2

Luria, A. R. (1976). *The Cognitive Development: Its Cultural and Social Foundations. (Trans. M. Lopez-Morillas & L. Solotaroff).* London: Harvard University Press.

Maher, C. A., & Sigley, R. (2014). Task-Based Interviews in Mathematics Education. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp. 579–582). Dordrecht, The Netherlands: Springer. https://doi.org/10.1007/978-94-007-4978-8_147

Mariotti, M. A. (2000). Introduction to proof: the mediation of a dynamic software environment. *Educational Studies in Mathematics, 44*(1/2), 25–53. https://doi.org/10.1023/A:1012733122556

Mariotti, M. A. (2006). Proof and proving in mathematics education. In A. Gutiérrez, & P. Boero (Eds.), *Handbook of research on the Psychology of Mathematics Education: Past, present and future* (pp. 173–204). Rotterdam, The Netherland: Sense.

Mariotti, M. A. (2013). Introducing students to geometric theorems: How the teacher can exploit the semiotic potential of a DGS. *ZDM, 45*(3), 441–452. https://doi.org/10.1007/s11858-013-0495-5

Martí, M. A., Taulé, M., Nofre, M., Marsó, L., Martín-Valdivia, M. T., & Jiménez-Zafra S. M. (2016). La negación en español: análisis y tipología de patrones de negación. *Procesamiento del Lenguaje Natural*, *57* (pp. 41–48). Sociedad Española para el Procesamiento del Lenguaje Natural. Jaén, España.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41–51. https://doi.org/10.2307/749097

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics*, *15*(3), 277–289. https://doi.org/10.1007/BF00312078

Meissner, H. (1986). Cognitive conflicts in mathematics learning. *European Journal of Psychology of Education*, *1*(2), 7–15. https://doi.org/10.1007/BF03172566

Mejía-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, *79*(1), 3–18. https://doi.org/10.1007/s10649-011-9349-7

Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction*, *20*(1), 79–118. https://doi.org/10.1207/S1532690XCI2001_4

Morris, A. K. (2007). Factors affecting pre-service teachers' evaluations of the validity of students' mathematical arguments in classroom contexts. *Cognition and Instruction*, 25(4), 479–522. https://doi.org/10.1080/07370000701632405

Morselli, F. (2006). Use of examples in conjecturing and proving: An exploratory study. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 185–192). Prague: PME.

Müller, U., Carpendale, J., & Smith, L. (2009). Introduction: Overview. In U. Müller, J. Carpendale, & L. Smith (Eds.), *The Cambridge Companion to Piaget* (Cambridge Companions to Philosophy, pp. 1–44). Cambridge: Cambridge University Press. https://doi.org/10.1017/CCOL9780521898584

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Netz, R. (1999). *The shaping of deduction in Greek mathematics: A study in cognitive history*. Cambridge UK: Cambridge University Press.

Newstead, S. E., & Griggs, R. A. (1983). Drawing inferences from quantified statements: A study of the square of opposition. *Journal of verbal learning and verbal behavior*, *22*(5), 535–546. https://doi.org/10.1016/S0022-5371(83)90328-6

Ngicho, D. O., Karuku, S., & King'endo, M. (2020). Manifestations and meanings of cognitive conflict among mathematics students in Embu, Kenya. *Educational Research and Reviews*, *15*(11), 690–699. https://doi.org/10.5897/ERR2020.4061

Okazaki, M., & Koyama, M. (2005). Characteristics of 5th graders' logical development through learning division with decimals. *Educational Studies in Mathematics*, *60*(2), 217–251. https://doi.org/10.1007/s10649-005-8123-0

Ordoñez-Montañez, C. (2014). La construcción de la noción de división y divisibilidad de números naturales, mediada por justificaciones, en alumnos de tercer grado de nivel primaria [Tesis de Maestría en Educación Matemática, Pontificia Universidad Católica del Perú]. Lima, Perú.

Organisation for Economic Co-operation and Development [OECD] (2009). *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. OECD.

Paenza, A. (2010). *Matemática… ¿estás ahí? Sobre números, personajes, problemas y curiosidades*. Siglo Veintiuno Editores. Buenos Aires, Argentina.

Pasztor, A., & Alacaci, C. (2005). On people's incorrect either-or patterns in negating quantified statements: A study. In S. Brown, G. Karakok, K. H. Roh, & M. Oehrtman (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (Vol. 27, pp. 1714–1719). Hillsdale: Erlbaum.

Peled, I., & Zaslavsky, O. (1997). Counter-Examples That (Only) Prove and Counter-Examples That (Also) Explain. *FOCUS on Learning Problems in mathematics*, *19*(3), 49–61.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16–25. https://doi.org/10.2307/1176006

Piaget, J. (1977). Problems of equilibration. In M. H. Appel, & L. S. Goldberg (Eds.), *Topics in Cognitive Development Equilibration: Theory, Research, and Application* (pp. 3–13). Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-4175-8_1

Polya, G. (1945). *How to solve it.* Princeton, New Jersey: Princeton University Press.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. http://dx.doi.org/10.1002/sce.3730660207

Potari, D., Zachariades, T., & Zaslavsky, O. (2010). Mathematics teachers' reasoning for refuting students' invalid claims. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of the Sixth Congress of the European Society for Research in Mathematics Education* (CERME 6) (WG2, pp. 281–290). Lyon, France.

Prediger, S. (2019). Theorizing in design research: Methodological reflections on developing and connecting theory elements for language-responsive mathematics classrooms. *AIEM - Avances de Investigación en Educación Matemática, 15*, 5–27. https://doi.org/10.35763/aiem.v0i15.265

Real Academia Española & Asociación de Academias de la Lengua Española (2010). *Nueva gramática de la lengua española. Manual*. Madrid: España.

Recio, A. (2002). La demostración en Matemática. Una aproximación epistemológica y didáctica. In M. F. Moreno, F. Gil, M. Socas, & J. D. Godino (Eds.), *Actas del V Simposio de la Sociedad Española de Investigación en Educación Matemática*, 27–43. Universidad de Almería.

Reeves, T. (2006). Design research from a technology perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 52–66). London: Routledge.

Reid, D. A. (2002). Conjectures and refutations in grade 5 mathematics. *Journal for Research in Mathematics Education, 33*(1), 5–29. https://doi.org/10.2307/749867

Reid, D. A. & Knipping, C. (2010). *Proof in Mathematics Education. Research, Learning and Teaching*. Rotterdam: Sense.

Reid, D. A. (2011). Understanding proof and transforming teaching. In Wiest, L., & Lamberg, T. (Eds.), *Proceedings of the Thirty-third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 15–30). Reno NV: University of Nevada.

Reid, D. A., & Vallejo-Vargas, E. A. (2017). Proof-based teaching as a basis for understanding why. In T. Dooley, & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education* (CERME 10) (WG1, pp. 235–242). Dublin, Ireland: DCU Institute of Education and ERME. https://hal.science/hal-01865651

Reid, D. A., & Vallejo-Vargas, E. A. (2018). When is a generic argument a proof? In A. J. Stylianides, & G. Harel (Eds.), *Advances in Mathematics Education Research on Proof and Proving* (ICME-13 Monographs, pp. 239–251). Springer, Cham. https://doi.org/10.1007/978-3-319-70996-3_17

Reid, D. A., & Vallejo-Vargas, E. A. (2019). Evidence and argument in a proof based teaching theory. *ZDM*, *51*(5), 807–823. https://doi.org/10.1007/s11858-019-01027-x

Reiss, K. M., Hellmich, F., & Reiss, M. (2002). Reasoning and proof in geometry: Prerequisites of knowledge acquisition in secondary school students. In A. D. Cockburn, & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 113–120). University of East Anglia.

Revlin, R., & Leirer, V. O. (1980). Understanding quantified categorical expressions. *Memory & Cognition*, *8*(5), 447–458. https://doi.org/10.3758/BF03211141

Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 23*(2), 139–159. https://doi.org/10.1080/10573560601158461

Schoenfeld, A. H. (2011). Series editor's foreword: The soul of mathematics. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. xii–xvi). New York, NY: Routledge.

Scribner, S. (1975). Recall of classical syllogisms: A cross-cultural interpretation of errors on logical problems. In R. J. Falmagne (Ed.), *Reasoning, Representation and Process* (pp. 153–173). Hillsdale, NJ: Erlbaum.

Segal, J. (1999). Learning about mathematical proof: Conviction and validity. *The Journal of Mathematical Behavior*, *18*(2), 191–210. https://doi.org/10.1016/S0732-3123(99)00028-0

Segal, S. U. (2009). *Action research in Mathematics in education: A study of Master's programs for teachers* [Unpublished doctoral dissertation]. Montana State University. Retrieved from http://scholarworks.montana.edu/xmlui/bitstream/handle/1/2237/SegalS0509.pdf?sequence=1

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational studies in mathematics*, *29*(2), 123–151. https://doi.org/10.1007/BF01274210

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36. https://doi.org/10.2307/30034698

Sellers, M. E. (2018). When 'negation' impedes argumentation: The case of Dawn. *Proceedings of the 21st Annual Conference on Research in Undergraduate Mathematics Education* (pp. 242–256). San Diego, CA.

Sellers, M. E. (2020). *Students' Quantifications, Interpretations, and Negations of Complex Mathematical Statements from Calculus* [Doctoral Dissertation, Arizona State University].

Shinno, Y., Miyakawa, T., Mizoguchi, T., Hamanaka, H., & Kunimune, S. (2019). Some linguistic issues on the teaching of mathematical proof. In U. T. Jankvist,

M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (CERME 11) (pp. 318–325). Utrecht, the Netherlands: Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Sjøberg, S. (2010). Constructivism and Learning. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (Vol. 5, pp. 485–490). Oxford: Elsevier.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, construction, and appreciation. *Canadian Journal of Science, Mathematics and Technology Education, 3*(2), 251–267.

Steffe, L. P. (1990). Inconsistencies and Cognitive Conflict: A Constructivist View. *Focus on Learning Problems in Mathematics*, *12*, 99–109.

Steffe, L. P. (1991). The constructivist teaching experiment: Illustrations and implications. In E. von Glasersfeld (Ed.), *Radical Constructivism in Mathematics Education* (pp. 177–194). Dordrecht, The Netherlands: Kluwer.

Steffe, L. P., Thompson, P. W., & Glasersfeld, E. v. (2000). Teaching Experiment Methodology: Underlying Principles and Essential Elements. In A. E. Kelly, & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 267–306). Mahwah, NJ: Lawrence Erlbaum Associates.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*(3), 289–321. http://www.jstor.org/stable/30034869

Stylianides, G. J. (2008). An Analytic Framework of Reasoning-and-Proving. *For the Learning of Mathematics, 28*(1), 9–16. http://www.jstor.org/stable/40248592

Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education*, *11*(4), 307–332. https://doi.org/10.1007/s10857-008-9077-9

Stylianides, G. J., & Stylianides, A. J. (2008). Proof in School Mathematics: Insights from psychological research into students' ability for deductive reasoning. *Mathematical Thinking and Learning, 10*(2)*,*103–133. https://doi.org/10.1080/10986060701854425

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, *72*(2), 237–253. https://doi.org/10.1007/s10649-009-9191-3

Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, *40*(3), 314–352. http://www.jstor.org/stable/40539339

Stylianides, A. J. (2011). Towards a comprehensive knowledge package for teaching proof: A focus on the misconception that empirical arguments are proofs. *Pythagoras*, *32*(1), 1–10. https://doi.org/10.4102/pythagoras.v32i1.14

Stylianides, G. J., Stylianides, A. J., & Shilling-Traina, L. N. (2013). Prospective Teachers' Challenges in Teaching Reasoning-and-Proving. *International Journal of Science and Mathematics Education*, *11*(6), 1463–1490. https://doi.org/10.1007/s10763-013-9409-9

Stylianides, A. J., & Stylianides, G. J. (2014). Viewing "mathematics for teaching" as a form of applied mathematics: Implications for the mathematical preparation of teachers. *Notices of the American Mathematical Society, 61*(3), 266–276.

Stylianides, G. J., & Stylianides, A. J. (2014). The Role of Instructional Engineering in Reducing the Uncertainties of Ambitious Teaching. *Cognition and Instruction*, *32*(4), 374–415. http://www.jstor.org/stable/43941151

Stylianides, G. J. (2016). *Curricular resources and classroom use: The case of mathematics*. New York, NY: Oxford University Press.

Stylianides, G. J., & Stylianides, A. J. (2017). Research-based interventions in the area of proof: the past, the present, and the future. *Educational Studies in Mathematics*, *96*(2), 119–127. https://doi.org/10.1007/s10649-017-9782-3

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for Research in Mathematics Education* (pp. 237–266). Reston, VA: National Council of Teachers of Mathematics.

Stylianides, A. J. (2019). Secondary students' proof constructions in mathematics: The role of written versus oral mode of argument representation. *Review of Education*, *7*(1), 156–182. https://doi.org/10.1002/rev3.3157

Stylianides, G. J., & Stylianides, A. J. (2020). Posing new researchable questions as a dynamic process in educational research. *International Journal of Science and Mathematics Education*, *18*(1), 83–98. https://doi.org/10.1007/s10763-020-10067-9

Swan, M. (2001). Dealing with misconceptions in mathematics. In P. Gates (Ed.), *Issues in Mathematics Teaching* (pp. 147–165). London: Routledge Falmer Press.

Tabach, M.; Levenson, E.; Barkai, R.; Tsamir, P.; Tirosh, D. and Dreyfus, T. (2009). Teachers' knowledge of students' correct and incorrect proof constructions. In F.L. Lin, F.J. Hsieh, G. Hanna, & M. de Villiers (Eds.), *Proceedings of ICMI study 19 Conference: Proof and Proving in Mathematics Education* (Vol. 2, pp. 214–219). Taipei, Taiwan: The Department of Mathematics, National Taiwan Normal University.

Tabach, M., Barkai, R., Tsamir, P., Tirosh, D., Dreyfus, T., & Levenson, E. (2010a). Verbal Justification—Is it a proof? Secondary School Teachers' Perceptions. *International Journal of Science and Mathematics Education*, *8*(6), 1071–1090. https://doi.org/10.1007/s10763-010-9230-7

Tabach, M., Levenson, E., Barkai, R., Tirosh, D., Tsamir, P., & Dreyfus, T. (2010b). Secondary school teachers' awareness of numerical examples as proof. *Research in Mathematics Education*, *12*(2), 117–131. https://doi.org/10.1080/14794802.2010.496973

Tall, D. (1977). Cognitive conflict and the learning of mathematics. In *Proceedings of the First Conference of The International Group for the Psychology of Mathematics Education. Utrecht, Netherlands: PME.*

Tirosh, D., & Graeber, A. O. (1990). Evoking cognitive conflict to explore preservice teachers' thinking about division. *Journal for Research in Mathematics Education*, *21*(2), 98–108. https://doi.org/10.2307/749137

Tirosh, C., & Vinner, S. (2004). Prospective teachers' knowledge of existence theorems. In M. J. Høines, & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, p. 360). Bergen University College, Norway: PME.

Tsamir, P., & Tirosh, D. (1999). Consistency and representations: The case of actual infinity. *Journal for Research in Mathematics Education*, *30*(2), 213–219. https://doi.org/10.2307/749611

Tsamir, P., Tirosh, D., Dreyfus, T., Barkai, R., & Tabach, M. (2009). Should proof be minimal? Ms T's evaluation of secondary school students' proofs. *The Journal of Mathematical Behavior*, *28*(1), 58–67. https://doi.org/10.1016/j.jmathb.2009.04.002

Valiña, M. D., & Martín, M. (2016). The influence of semantic and pragmatic factors in Wason's selection task: State of the art. *Psychology*, *7*(6), 925–940. http://dx.doi.org/10.4236/psych.2016.76094

Vallejo-Vargas, E. A. (2012). *Análisis y propuesta en torno a las justificaciones en la enseñanza de la divisibilidad en el primer grado de secundaria* [Tesis de Maestría, Pontificia Universidad Católica del Perú]. Lima, Perú.

Vallejo-Vargas, E. A. & Ordoñez-Montañez, C. (2014). 3rd graders construction of division and divisibility knowledge only depending on a Key Notion. In Oesterle, S., Nicol, C., Liljedahl, P. & Allan, D. (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 6, p. 400). Vancouver, Canada: PME.

Vallejo-Vargas, E., A. & Ordoñez-Montañez, C. (2015). An example of proof-based teaching: 3rd graders constructing knowledge by proving. In Krainer, & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 230–231). Prague: Charles University and ERME.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Introduction to educational design research. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 3–7). London: Routledge.

Veel, R. (1999). Language, knowledge and authority in school mathematics. In F. Christie (Ed.), *Pedagogy and the shaping of consciousness: Linguistic and social processes* (pp. 185–216). London, England: Cassell.

Vender, M., & Delfitto, D. (2010). Towards a pragmatics of negation: The interpretation of negative sentences in developmental dyslexia. *GG@ G–Generative Grammar at Geneva*, *6*, 1–28.

Vinsonhaler, R., & Lynch, A. G. (2020). Students' Understanding of Counterexamples, *Mathematics Teacher: Learning and Teaching PK-12 MTLT*, *113*(3), 223-228. https://doi.org/10.5951/MTLT.2019.0092

Vinner, S. (1983). The notion of proof – some aspects of students' views at the senior high level. In R. Hershkowitz (Ed.), *Proceedings of the Seventh International Conference of the Psychology of Mathematics Education* (pp. 289–294). Rehovot, Israel: Weizmann Institute of Science.

Wason, P. C. (1966). Reasoning. In Foss, B. (Ed.), *New Horizons in Psychology* (pp. 135–151). Harmondsworth: Penguin Books.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly journal of experimental psychology*, *20*(3), 273–281. https://doi.org/10.1080/14640746808400161

Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah, NJ: Lawrence Erlbaum.

Waxer, M., & Morton, J. B. (2012). Cognitive Conflict and Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 585–587). https://doi.org/10.1007/978-1-4419-1428-6_280

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101–119. https://doi.org/10.1023/A:1015535614355

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, *12*(4), 306–336. https://doi.org/10.1080/10986065.2010.495468

Weber, K. (2014). Proof as a Cluster Concept. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the 38th North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 5, pp. 353–360). Vancouver: PME

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, *18*(4), 451. https://doi.org/10.1037/h0060520

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in mathematics education*, *27*(4), 458–477. https://doi.org/10.5951/jresematheduc.27.4.0458

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A Research Companion to Principles and Standards for School Mathematics* (pp. 22–44). Reston, VA: National Council of Teachers of Mathematics.

Yang, K.-L., & Lin, F.-L. (2008). A model of reading comprehension of geometry proof. *Educational Studies in Mathematics, 67*(1), 59–76. https://doi.org/10.1007/s10649-007-9080-6

Yopp, D. A. (2017). Eliminating counterexamples: A Grade 8 student's learning trajectory for contrapositive proving. *The Journal of Mathematical Behavior*, *45*, 150–166. https://doi.org/10.1016/j.jmathb.2017.01.003

Yu, J. Y. W., Chin, E. T., & Lin, C. J. (2004). Taiwanese junior high school students' understanding about the validity of conditional statements. *International Journal of Science and Mathematics Education*, *2*, 257-285.

Zaslavsky, O., & Peled, I. (1996). Inhibiting factors in generating examples by mathematics teachers and student teachers: The case of binary operation. *Journal for research in Mathematics Education*, *27*(1), 67–78. https://doi.org/10.2307/749198

Zaslavsky, O., & Ron, G. (1998). Students' understandings of the role of counterexamples. In A. Olivier, & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 225–232). Stellenbosch, South Africa.

Zaslavsky, O. (2005). Seizing the Opportunity to Create Uncertainty in Learning Mathematics. *Educational Studies in Mathematics*, *60*(3), 297–321. https://doi.org/10.1007/s10649-005-0606-5

Zaslavsky, O., & Shir, K. (2005). Students' conceptions of a mathematical definition. *Journal for Research in Mathematics Education*, *36*(4), 317-346. https://doi.org/10.2307/30035043

Zaslavsky, O. (2008). Meeting the challenges of mathematics teacher education through design and use of tasks that facilitate teacher learning. In B. Jaworski, & T. Wood (Eds.), *The International Handbook of Mathematics Teacher Education: The Mathematics Teacher Educator as a Developing Professional* (Vol. 4, pp. 93–114). The Netherlands: Sense Publishers.

Zaslavsky, O., Nickerson, S. D., Stylianides, A. J., Kidron, I., & Winicki-Landman, G. (2012). The need for proof and proving: Mathematical and pedagogical perspectives. In G. Hanna, & M. de Villiers (Eds.), *Proof and Proving in Mathematics Education: the 19th ICMI Study* (pp. 169–190). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2129-6_9

Zazkis R. & Chernoff E. (2006). Cognitive conflict and its resolution via pivotal/bridging example. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th Annual Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 465–472). Prague, Czech Republic: PME.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics*, *68*(3), 195–208. https://doi.org/10.1007/s10649-007-9110-4

Zazkis, R., Liljedahl, P., & Chernoff, E. J. (2008). The role of examples in forming and refuting generalizations. *ZDM*, *40*(1), 131–141. https://doi.org/10.1007/s11858-007-0065-9

Appendix

Appendix Notations: Notations used in dialogues and quotations

Notations used in dialogues and quotations

- "[---]" indicates intelligible parts in the speech.
- "…" indicates parts omitted that do not change the purpose of the parts that have been included.
- "[X]" includes clarifications "X" that I completed, which could be inferred from what has been said.
- "(X)" includes clarifications or details for gestures or voice tones used that are not said but might be important to be aware of.
- A *hyphen* (–) at the end of a (complete or incomplete) word indicates interruptions in the speech.
- The use of words in *capital letters* indicate that the speaker emphasized such words for example by raising her/his voice
- "( )" means basically a few seconds without saying anything. It might be because either the person takes some time to think on what s/he is going to say next, or to indicate s/he does not finish the sentence with the intention that another person continues the sentence.

Appendix EI1-P1: Exploratory Interview 1 - Part 1

Exploratory Interview 1 - Part 1

All the questions that follow should be answered from your personal perspective, your own ideas. There is no attempt to evaluate you in this process.

1) About Mathematics in general

(1.1) From your point of view:

a) What is Mathematics? (when you think about Mathematics, what is it the first thing that comes to your mind?)
b) What does it mean *"to do mathematics"*?

(1.2) Do you think Mathematics is really useful for daily life as it is typically claimed?

- If she says YES → Point out in what aspects for example.
- If she says NO → Why not?

2) About the teaching and learning of Mathematics

(2.1) What's your *main goal* when you *teach mathematics*?

(2.2) How do you know/test that your students *understand* what you address in the Mathematics classroom?

(2.3) Do you foster *mathematical reasoning* (MR) in your Math lessons?

- If she says YES:
  → What is your goal when doing so?
  → Give some examples of how you promote MR.
- If she says NO → Why not?

3) About Mathematical Proof

(3.1) Observe (show them) the following terms:

→ Proof
→ Justification
→ Explanation
→ Verification

a) Do you use some of these terms in your Math lessons?
   - If she says YES → Point which one and give an example of how you use them.
b) Do you find any differences among them?
   - If she says YES → Provide some examples
c) Do you make any explicit distinction of these terms in your classes?
   - If she says YES → Elaborate

(3.2) Do you know/remember any mathematical property/theorem and its proof?

- If she says YES → Provide an example
- If she says NO → Show an example (Example 1: the sum of two odd numbers is an even number; Example 2: The product of two number is commutative) and ask if it is a property and how she knows it.

(3.3) Do you think it's possible that an elementary school student can prove mathematical properties?

- If she says YES → What age do you think is the best age to start from? Can you give an example (property and age)?
- If she says NO → Why not?

(3.4) Do you think it's important that your students prove properties in their classrooms?

- If she says YES → Provide examples of properties they could prove and explain why this would be important.

4) About the mathematical content: Division

(4.1) How do you usually teach this topic?

(4.2) What is your main goal when introducing this topic to third graders?

(4.3) Use Appendix EI1-P2.

<u>Focus on</u>: Section I only.

(Work "together". Verify that the problems posed are clear enough before the teacher answers the questions)

(4.4) Do you know what the relationship between the divisor and the remainder in any division of natural numbers is? (i.e. what values can take the remainder according to the divisor's value?)

- If she says YES → which one?

(4.5) Are third graders taught this relationship?

- If she says YES → How is it taught? How is it introduced?
- If she says NO → When is it taught then? Grade?

(4.6) Do you understand why the remainder must be less than the divisor?

- If she says YES → explain why and how would you explain this to a third grader?

Appendix EI1-P2: Exploratory Interview 1 - Part 2

Exploratory Interview 1 - Part 2

Name:

Section I

First, consider the following remarks when solving the problems posed:

- There is no need to solve the divisions with *decimals* or *fractions*. All the following divisions are divisions with natural numbers.
- Consider 0, 1, 2, 3, and so on, as the natural numbers.
- Remember what are the elements of any division:

Dividend	Divisor
(Remainder)	Quotient

1) Solve the following division:

Divide 22110011 by 11

a) The quotient is: ___________________

b) The Remainder is: ____________________

2) Now, think only of divisions by 3; it means, think of divisions with the divisor equal to 3 in all cases.

a)	b)	c)
Write down an example of division with the given characteristics.	Solve the division you provided in a) and point to the remainder value with an arrow.	Is it possible to determine all the values the remainder can take when dividing by 3 (when the divisor is equal to 3)? ▪ If your answer is YES, provide all these values. ▪ If your answer is NO, explain why.

3) Is 1 divisible by 5? Why?

4) Give two examples of numbers divisible by 6.

Choose one of your examples and explain why this is divisible by 6.

Example 1:	Explanation:
Example 2:	

Section II

1) Explain why the following property is true:

"The sum of two even numbers is an even number"

2) Given the statement: "All VALLEJO numbers are even numbers"

From the following list of statements, select those that convey/communicate the same as the given statement (above).

a) "Some VALLEJO numbers are even numbers"
b) "All even numbers are VALLEJO numbers"
c) "If we have a VALLEJO number then it is an even number"
d) "No VALLEJO number is an even number"
e) "The even numbers are not VALLEJO numbers"
f) "Some VALLEJO numbers are not even numbers"
g) "Every VALLEJO number is an even number"
h) "The VALLEJO numbers are even numbers"
i) "If a number is even then it is a VALLEJO number"
j) "There are VALLEJO numbers that are even numbers"
k) "There are no VALLEJO numbers that are not even"
l) "Not all VALLEJO numbers are even numbers"

3) Read the following classroom episode and then answer the given questions.

In a Mathematics class, one of the students, Rodrigo, raises his hand to share with the whole class one of his most recent “mathematical discoveries”:

Rodrigo	*If we multiply any natural number by 5, we will get a number with the units-digit equal to 0 or 5*
Sebastian (another student in the same class)	*That’s true because if you multiply 11 by 5, you get 55, which ends in 5. If you multiply 24 by 5, the result is 120, which ends in 0.* *So, yes, you will always get what you just said*

According to this classroom episode:

a) What score would you give to Sebastian’s justification for Rodrigo’s “mathematical discovery”?

   Think of scores from 1 to 5, where 5 is the best score. Explain why you gave this score and not a different one.

b) Do you think that Sebastian’s justification shows/guarantees that Rodrigo’s “mathematical discovery” is true? Why? How do you know?

Appendix CE8: Classroom Episode 8

Classroom Episode 8.1

Context: a third-grade class is discussing the case of fair, whole and maximal (FWM) distributions between two people.

Teacher	If I follow the three conditions, can I have one object left?
María	Yes.
Teacher	Let's see, give me an example where there is one object left, between two people.
María	Nine objects among two people, there would be one object left
Teacher	Then can I have one object left?
Students	Yes!!!
Teacher	There might be one object left, indeed, good. What other number of objects left can I have left?
Carlos	Three.
Teacher	Can I have three objects left if I distribute among two people?

Classroom Episode 8.2

Context: the previous discussion continues, now with a focus on FWM distributions among 3 people.

Teacher	Can I have 5 objects left?
	(…)
Teacher	Can I have 4 objects left?
	(…)
Teacher	And can I have 3 objects left? That I can have left, right?
	(…)
Erica	You can only have left number 2 and number 1
Teacher	Could you repeat what you just said?
Erica	For the number 3, you can only have left 2 and 1, because they are smaller numbers. If there are 4 people, there might be left 3, 2, or 1.
Teacher	Why can't we have more left?
Erica	If there were more objects than the number of people, you can continue distributing objects. If there were fewer objects than the number of people, the distribution is over.
Teacher	What do you think? Does it only work when I distribute objects among 3 people? When else?
Olinda	With 2 people
Teacher	Does it also work with 2 people?
Felipe	With 4
Teacher	Does it also work with 4 people?
Pepe	In the case of 2, we couldn't have 2 left, but only 1 or 0
Teacher	Aha! I forgot about 0. When we have distributions among 3 people, then how many objects can we have left?
Students	0, 1, 2
Teacher	What else?
	Some students answered: 3; other students refuted that answer saying: no!

Classroom Episode 8.3

Context: the class is analyzing some students' responses to one of the activities they solved. Precisely, the class is discussing the following task:

When we distribute in a fair, whole and maximal way among **3** people, is it possible that there are 5 objects left once the distribution is over?

Yes   No

Why?

Students	No, because it is possible to continue the distribution…
Teacher	… and if it is possible to continue the distribution, what condition does it hold? Let's see, hat condition does it hold Alexandra? If we can continue with the distribution…
Alexandra	Fair, whole and maximal.
Teacher	Fair, whole and maximal, right.

Appendix CE9: Classroom Episode 9

Classroom Episode 9

Context: a group of teachers is analyzing the truth value of the following statement:

"The numbers divisible by 6 are divisible by 3"

<table>
<tr><td>María</td><td>It is false</td><td></td></tr>
<tr><td>Erica</td><td>Why do you claim that it is false? How do you know?</td><td></td></tr>
<tr><td>María</td><td>I did the following (María shows what she has done in her notebook, see the following table)</td><td></td></tr>
<tr><td></td><td>
<table>
<tr><th>Numbers that are divisible by 6</th><th>Numbers that are divisible by 3</th></tr>
<tr><td>6
12
18
24
30
…</td><td>3
6
9
12
15
18
21
…</td></tr>
</table>
</td><td></td></tr>
<tr><td>María</td><td>And I realized that 9 is divisible by 3, but it is not divisible by. There is no number 9 in my list of numbers that are divisible by 6. That is why I say that this is a false statement.</td><td></td></tr>
</table>

Appendix CE13: Classroom Episode 13

Classroom Episode 13.1

Context: the class is analyzing the students' answers for one of the activities about the relationship between the number of people to whom the objects are shared out and the number of objects that might be left once the fair, whole and maximal (FWM) distribution is finished. Namely, the following problem is revisited:

When fair, whole and maximal distributions are performed among **3** people, is it possible that 5 objects are left once the distribution among the 3 people is finished?

Yes No

Why?

Pedro	But miss, miss, over there you are not told how many objects there are at the beginning
Teacher	Is it necessary to know how many objects are distributed at the beginning? Andy?
Andy	Over there it says, is it possible that 5 objects are left? It doesn't say that there were 5 objects at the beginning
Teacher	It doesn't say that the objects you are given 5 objects at the beginning. It says, if there were 5 objects left, is it possible to distribute them among 3 people, yes or no?

Classroom Episode 13.2

Context: the class is discussing the truth value of the statement given by Carlita; that is:

"If a distribution is FWM, then the remainder is 0"

The students were expected to justify their answer.

Teacher	What Carlita says, that the whole and fair distributions, if they end up with a remainder zero, then they also hold the maximal condition.
Vale (and some other students)	Yes, yes, that's true.
Teacher	Is that true?
Some students	No…
Teacher	Is that true?
Students	Yes!!!
Teacher	Is Carlita's statement true?
	Some students answer "yes", but do not sound very confident
Teacher	Is this zero (the teacher points out the remainder in the division 210 by 9 on the whiteboard)? Is this zero (the teacher points out the remainder in the division 6 by 5 on the whiteboard)?
Students	No.
Teacher	Then, is it true what Carlita says?
Students	No!
Teacher	No, and we verified that here. Is that clear? (the teacher points to the examples on the whiteboard) Do you understand? Then, what can I say? How can I conclude this here?

Appendix A5: Activity 5

Activity 5 / Task 1

Name:

Carlitos is playing to distribute objects in a fair, whole and maximum way. Carlitos is only working with distributions among **3 people**.

Carlitos says:

"I'm already done with my distribution and I have ***5*** *objects left"*

Now you answer the following question:

Is it possible that what Carlitos says is true?

Yes	No

- If you marked "Yes", write down an example of a FWM-distribution among 3 people so that you have 5 objects left once you finished your distribution.
- If you marked "No", explain clearly why Carlitos cannot have 5 objects left once the distribution is over.

Activity 5 / Task 2

Name:

a) In a division by 4, what does number 4 represent?

b) What does it mean to divide by 4?

c) Is it possible that in a division by 4 the remainder is equal to 6? Why?

Appendix A6: Activity 6

Activity 6

Members of the group:

- ____________________________________
- ____________________________________

Answer each of the following tasks:

1) What does it mean "to divide by 3" according to the model of distributions?

2) Complete the following table with suitable numbers, accordingly.

When you divide by…	**The remainder could be equal to… [write down all possibilities]**	**The MINIMAL remainder is… ("*minimal*" means the smallest value)**	**The MAXIMAL remainder is… ("*maximal*" means the highest value)**
2			
3			
10			
17			

Continue

3) Complete the following sentences with a suitable number respectively so that every sentence once completed is a true sentence.

a) When it is a division by 23, the **maximal remainder** is ________

b) When it is a division by 58, the **maximal remainder** is ________

c) From sentence (b), explain WHY the remainder cannot be a number bigger than the one you wrote down in the blank space.

Appendix A10: Activity 10

Activity 10 / Task 1

Name:

Observe the following box.

In this box I have hidden certain number of marbles. I will not tell you how many marbles I put in there; however, I can tell you that if I would distribute these marbles among **6** people in a fair, whole and maximal way, there would be 0 marbles left.

Now, answer the following questions:

a) In spite of the fact that the number of marbles in the box is not specified in the formulation of the task, it is possible to imagine what possibilities there are for it. What could be the number of marbles inside the box? Suggest at least two examples.

____

b) If I would distribute the marbles in the box among 3 people in a fair, whole and maximal way, would there be 0 marbles left as well?

Yes	No

Why?

Activity 10 / Task 2

Name:

Imagine that you have a magic bag. Inside this magic bag you can find all numbers divisible by 4.
Now imagine that you take away, one by one, the numbers from the magic bag.

Answer the following questions:

a) Are some of the numbers that you take from the magic bag divisible by 8?

Yes No

Why?

b) Are all the numbers that you take from the magic bag divisible by 8?

Yes No

Why?

Appendix A12: Activity 12

Activity 12 / Task 1

Name:

Given the following structure:

16 is divisible by ________

a) Choose a number and use it to complete the given structure above in such a way that the complete sentence is true.

16 is divisible by ________

Once you complete the sentence, justify why it is true. Explain the meaning of each number you use in the sentence according to the fair, whole and maximal distributions model.

b) Write down all the possible true sentences using the given structure.

____

____

____

____

____

____

Activity 12 / Task 2

Name:

Julito is a curious student. He solved Task 1 not only with number 16, but with other numbers. Observe the following table that summarizes the order and the examples Julito tried out $(3, 6, 16, 24)$.

3 is divisible by 1	**6** is divisible by 1	**16** is divisible by ___	**24** is divisible by 1
3 is divisible by 3	**6** is divisible by 2	**16** is divisible by ___	**24** is divisible by 2
	6 is divisible by 3	**16** is divisible by ___	**24** is divisible by 3
	6 is divisible by 6	**16** is divisible by ___	**24** is divisible by 4
		16 is divisible by ___	**24** is divisible by 6
			24 is divisible by 8
			24 is divisible by 12
			24 is divisible by 24

After finishing his exploration with number 24 (last column in the table), Julito says to his teacher:

I have a conjecture! My conjecture is:

"Every time we work with a bigger number, we will always obtain more true sentences than in the previous cases."

Now you answer the following question:

Is Julito's conjecture true or false? Why?

Appendix A13: Activity 13

Activity 13

Instructions:

(Give the following statements:)

Statement 1: *Not all divisions of natural numbers give a remainder 0*	**Statement 2:** *All divisions of natural numbers do not give a remainder 0*
Statement 3: *All divisions of natural numbers give a remainder 0*	**Statement 4:** *There exist divisions of natural numbers that give a remainder 0*
Statement 5: *There do not exist divisions of natural numbers that give a remainder 0*	**Statement 6:** *There exist divisions of natural numbers that do not give a remainder 0*
Statement 7: *No division of natural numbers gives a remainder 0*	**Statement 8:** *There do not exist divisions of natural numbers that do not give a remainder 0*

1) Group the given statements in such a way that each group include statements that state exactly the same thing.

2) Answer the following questions:

   2.1) Are the statements universal or existential statements?

   2.2) Are they true or false?

Appendix A14: Activity 14

Activity 14

Members of the group:

- ____
- ____

1) Every member of the group writes down a number that is divisible by 1. The chosen numbers must be different from each other.

Example 1:	Example 2:	Example 3:

2) How would you explain that the numbers you chose are divisible by 1? Pick one of your examples and justify why this number is divisible by 1.

3) Is possible to give more examples of numbers that are divisible by 1?

Yes   No

- If your answer is "Yes", provide three additional examples.
- If your answer is "No", explain why this is not possible.

Example 4:	Example 5:	Example 6:

4) Is possible to write down all the numbers that are divisible by 1?

Yes | No

- If your answer is "Yes", write down all these numbers.
- If your answer is "No", explain why this is not possible.

5) What characteristics do the numbers that are divisible by 1 have?

6) Justify why all these numbers satisfy the condition of being divisible by 1.

Appendix A15: Activity 15

Activity 15

Name:

1) Complete the following sentences with suitable numbers so that once complete the sentence is true.

_________ **is divisible by 4**
_________ **is divisible by 4**
_________ **is divisible by 4**
_________ **is divisible by 4**

Your examples should not be repeated

2) Write down one more sentence with the same structure from above:

____

3) Is possible to have more true sentences with the same structure?

Yes | No

How many in total?

____

4) How many numbers are divisible by 4?

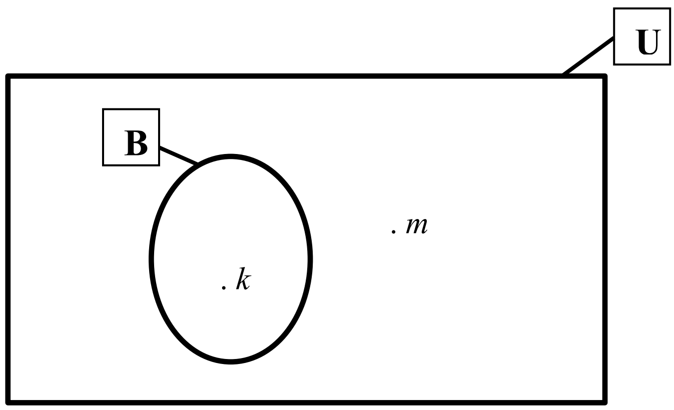
____

Explain why.

Appendix EA1: Extra Activity 1

Extra Activity 1

1) Statement: *“All natural numbers bigger than 5 are RAINBOW numbers”*
   1.1) Let’s suppose that this statement is true. What kind of mathematical evidence would be sufficient to show in order to guarantee that it is true?
   1.2) Which is the set of analysis?
   1.3) How many elements does the set of analysis have?
   1.4) In the given statement, is it asserted that the RAINBOW numbers are bigger than 5?
   1.5) In the given statement, is it asserted that some natural numbers bigger than 5 are RAINBOW?
2) Statement: *“All Peruvians are BLUMEN”*
   2.1) To justify that the statement is false, what kind of mathematical evidence would be sufficient to show?
   2.2) What specific characteristics should have this evidence?
   2.3) Which is the set of analysis?
   2.4) Does the set of analysis have infinite elements?
   2.5) In the given statement, is it asserted that at least one Peruvian is *BLUMEN*?
3) Statement: *“Some RAINBOW numbers are smaller than 3”*
   3.1) What kind of mathematical evidence is sufficient to guarantee that the statement is true?
   3.2) Which is the set of analysis?
   3.3) In the given statement, is there anything asserted about all *RAINBOW* numbers? Explain.
   3.4) In the given statement, is it asserted that some *RAINBOW* numbers are not smaller than 3?
4) Given set B as the set of all *BEAUTIFUL* numbers and a representation for it (see below). U denotes the universal set that includes set B.



   Based on this representation, answer the following questions:

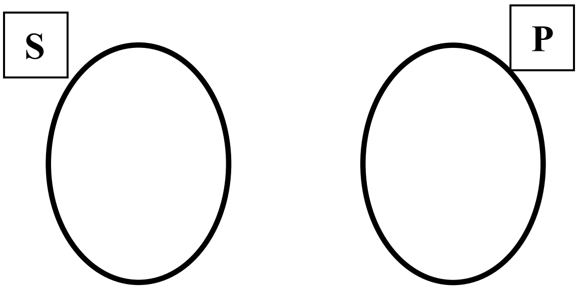
   4.1) What can you say about “*k*” with regard to the characteristics of the elements of set B?

4.2) What can you say about "*m*" with regard to the characteristics of the elements of set B?
4.3) In the given representation (above), locate "*p*" that is a number that is not *BEAUTIFUL.*
4.4) In the given representation, paint the region(s) where the numbers that are not *BEAUTIFUL* are located.

5) Given sets S and P,

S: the set of *SOBRINO* numbers
P: the set of *PERFECTO* numbers

The following representation shows the relation between sets S and P.



According to this information, answer the following questions:

5.1) In the given representation, draw an "*x*" where a number that is *PERFECTO* can be located.

5.2) In the given representation, paint the region(s) where the numbers that are not *PERFECTO* can be located.

5.3) With a "*z*" locate a number that is not *PERFECTO*.

5.4) Complete the following statements with "ARE" o "ARE NOT" so that the statement is true according to the representation:

a) The *PERFECTO* numbers ____________ *SOBRINOS* numbers.

b) All *SOBRINO* numbers ____________ *PERFECTO* numbers.

c) The *SOBRINO* numbers ___________ numbers that __________ *PERFECTO* numbers.

6) Statement: *"No PERFECTO number is BELLO"*

6.1) Use a Venn/Euler diagram to represent the statement. Write down in detail a description for the sets in the diagram.

6.2) According to the diagram, what can be asserted about the BELLO numbers?

6.3) Complete the following sentence so that it is a true sentence:

If a number is PERFECTO, then it ______________ *BELLO.*

6.4) If a number is not BELLO, is it PERFECTO?

Appendix EA2: Extra Activity 2

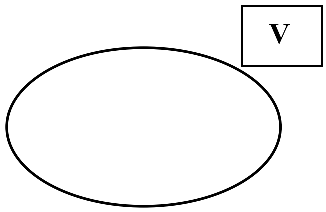
Extra Activity 2

1) Analyze the following statements:

   Statement 1: No even number is a *voilá* number.

   Statement 2: No even number is not a *voilá* number.

   a) Let's divide the whiteboard in three parts: the first part for the Statement 1; the second part for the statement 2; the third part for the statement 3, which comes below.
   b) Use a Venn or Euler diagram to represent Statement 1.
   c) The following is a diagram for the set V, the set of all *voilá* numbers:

   Where would be located the numbers that are NOT *voilá* numbers?

   d) How can Statement 2 be re-written in such a way that it keeps the following structure?

   Statement 3: No even number is a ________________________

   Are Statements 2 and 3 equivalent statements?

   e) What sets are involved in Statement 3?
   f) Use a Venn or Euler diagram to represent Statement 3.
   g) In the diagram for Statement 3, where are located the *voilá* numbers?
   h) How are sets E (set of even numbers) and V related in this statement?
   i) Write down the Statement that relates sets E and V, based on the diagram of the Statement 3.

2) Given the statement:

   Statement 4: There do not exist even numbers that are *voilá* numbers

   Answer the following question:

   Is Statement 4 equivalent to "No even number is *voilá*"? Why?

3) Given the statement:

   Statement 5: There do not exist natural numbers that cannot be divided by 2.

   Answer the following:

   Is Statement 5 equivalent to

   "No natural number can be divided by 2"?

Appendix D1: Tasks in Discussion 1 of the 2018-intervention

Sub-Discussion	Tasks in Discussion 1 of the 2018-intervention
Discussion 1.0	Pablito's statement: ***All divisions of natural numbers are exact divisions*** Now answer the following: 1) Is Pablito's statement true? 2) Each teacher individually justifies her answer on the whiteboard.
Discussion 1.1	Statement: ***All divisions of natural numbers are exact divisions*** We have justified that this is a false statement. See a justification. **In order to *justify* that the statement is false, could have it been sufficient to say or write down:** ***"It is false, because not all divisions of natural numbers are exact divisions"*? Why?**
Discussion 1.2	Statement: ***All divisions of natural numbers are exact divisions*** **Was it necessary that we all provide exactly the same *counterexample* as justification? Why?**
Discussion 1.3	Statement: ***All divisions of natural numbers are exact divisions*** **Is it sufficient to provide only one counterexample to justify that the *statement* given is false? How many otherwise?**
Discussion 1.4.1: Classroom Episode	**Context:** a third-grade class is analyzing the statement: ***"All divisions of natural numbers are exact divisions'***. The class suspects this is a false statement. The teacher asks his students to discuss the justification in groups. The teacher walks around the students' desks in order to listen to the discussions. - Leo: 5 divided by 1 - Irma: 20 divided by 8 - Teacher: and your example Ariana? - Ariana: 9 divided by 5 - Teacher: very good **Does the example given by Leo justify that the statement is false? Why?**
Discussion 1.4.2: Classroom Episode	Classroom Episode: [same classroom episode as in Discussion 1.4.1] **Does the example given by Leo is sufficient to justify that the statement is true? Why?**
Discussion 1.5: Another example	Statement: ***The people present in this classroom are minors*** a) In the statement, is it stated that ***some people present in this classroom are minors****?* Why? b) What does this statement exactly state? **Universal/General Statements**
Discussion 1.6.1	Statement: ***The people present in this classroom are minors*** a) Is this statement universal? How do you know? b) Is this statement true? c) Justify this **truth value** d) What characteristic(s) does the counterexample that justifies the statement is false have?
Discussion 1.6.2	Statement: ***All divisions of natural numbers are exact divisions*** a) What is exactly stated in this statement? b) We have seen this statement is false. Observe the Justification that proves this. What characteristic(s) does the counterexample that justifies the statement is false have? c) Is the statement given universal? How do you know?
Discussion 1.7: Classroom Episode 9	Statement in the episode: ***All numbers divisible by 6 are divisible by 3*** a) Is the case given by María (number 9) a counterexample for the statement given? What characteristics should have a counterexample that disprove the statement given? b) In the statement given, what is **the information given**? c) What should I analyze about this **information given**?

Appendix D2: Tasks in Discussion 2 of the 2018-intervention

Sub-Discussion	Tasks in Discussion 2 of the 2018-intervention
Discussion 2.1	Consider the following statement: ***If a distribution is FWM then the remainder is zero*** a) Is it ***true***? Why? b) In this statement, is it stated that ***some FWM distributions have a remainder zero***? c) What is stated in this statement? d) Is it ***equivalent*** to a ***universal*** statement? Which one? e) What do we need to show in order to guarantee that the statement is false? **Conditional Statement: *"If… then…"***
Discussion 2.2	Consider the following statement: ***If a distribution is FWM, then the number of objects left is less than the number of people*** a) What is stated in this statement? b) Is it ***true***? c) Is it equivalent to a ***universal statement***? d) Write down a universal statement that states the same thing as the statement given
Discussion 2.3 (Classroom Episode 9)	Statement in Classroom Episode 9: ***All numbers divisible by 6 are divisible by 3*** Write down the conditional statement that states exactly the same thing as the statement included in the episode does. Questions analyzed before: a) The case given by Maria (9), is it a counterexample for the statement given? b) In the statement given, what is the ***information given***? c) What should I analyze about this information given?
Discussion 2.4	**Classroom Episode 13 (see Appendix CE13)**
Discussion 2.4.1	Consider the following statement: ***If a distribution is FWM, then the remainder is zero*** Answer the following question: In this statement, is it stated that ***if the remainder of a distribution is zero, then the distribution is FWM***? Explain.
Discussion 2.4.2	Consider the following statement: ***If a number is even, then the number is FAST*** Answer the following question: In this statement, is it stated that ***if a number is FAST, then it is even***? Explain.
Discussion 2.4.3	Consider the following statement: ***If the units-digit of a number is 0, then the number is divisible by 10*** Now answer: In this statement, is it stated that ***if a number is divisible by 10, then the units-digit of the number is 0***? Explain.
Discussion 2.4.4	Consider the following statement: ***If the units-digit of a number is 0, then the number is divisible by 5*** Now answer: In this statement, is it stated that ***if a number is divisible by 5, then the units-digit of the number is 0***? Explain.
JV3 – Task 1	Alberto claims the following: ***"If a distribution is whole and fair, then it is FWM"*** ***Is Alberto's statement true? Why?***
JV3 – Task 2	Benito states what follows: ***"If a distribution is FWM, then it is fair"*** ***Is Benito's statement true? Why?***
JV3 – Task 3	Diego states that: ***"If a distribution is FWM, then it is whole"*** ***Is Diego's statement true? Why?***
JV3 – Task 4	Ernest states that: ***"If a whole and fair distribution has a remainder zero, then the distribution is FWM"*** ***Is Ernest's statement true? Why?***

Appendix D3: Tasks in Discussion 3 of the 2018-intervention

Sub-Discussion	Tasks in Discussion 3 of the 2018-intervention
Activity 12	(See Appendix A12)
Discussion (Post Activity 12)	Julito's conjecture: ***"Every time we work with a bigger number, we will always get more true sentences than in the previous cases"*** a) Is it true? b) Is it a universal statement? c) What kind of justification is sufficient to guarantee this truth value?
Discussion 3.1 (Discussion from slide 69 revisited)	When you divide **3** by **5**, what is the **result**? When you divide **4** by **8**, what is the **result**? And when you divide **3** by **4**, what is the **result**? And when you divide **2** by **6**? We formulated first the **CONJECTURE:** *"When we divide a number C by a number A bigger than C, the result is zero and the remainder is equal to C"* We provided the **JUSTIFICATION** Then, the conjecture became a **MATHEMATICAL TRUTH**
Discussion 3.2	Pepito has formulated the following conjecture: ***"All palindrome numbers are divisible by 11"*** Pepito claims his conjecture is true ***because*** he has verified it with "big" numbers such as: 121, 2442, 3553, 123321, which are divisible by 11. He used his calculator to support his work. **Is Pepito's conjecture true?** **Is his justification valid?** **Remember:** A palindrome number is the number that reads the same backwards.
Task	Given the following expression: $\mathbf{1 + 1141n^2}$ a) For $n = 1$, does the expression produce a perfect square? b) For $n = 2$, does the expression produce a perfect square? c) For $n = 3$, does the expression produce a perfect square? d) What do you observe? e) Can you formulate a conjecture based on your observations? f) Is your conjecture true? Why? **Remember:** A perfect square is a number that can be expressed as the square (power) of a natural number. For example: 25 is a perfect square since it can be written as 5^2; 1 is a perfect square; 121 is a perfect square; etc.
Discussion (Task)	**CONJECTURE:** For all n natural number (different from zero), the following expression: $1 + 1141n^2$ never produces a perfect square. Is the conjecture true or false? **CAREFUL! It is FALSE!** For n from **1** until **30,693,385,322,765,657,197,397,207** this expression does not produce a perfect square; however it does for the next natural number. Computers were used to verify this.

Discussion 3.3	In 1650, **Pierre de Fermat** conjectured that all natural numbers of the form $2^{2^n}+1$, with *n* a natural number, were **prime numbers**. In 1732, **Leonhard Euler** proved that this conjecture was false. He showed that when $n = 5$ the Fermat number is not a prime number $2^{2^5}+1 = 2^{32}+1 =$ 4294967297 = (641).(6700417)
Discussion 3.4	**Goldbach's** (1742) strong conjecture: ***"Every even number bigger than 2 can be written as the sum of two prime numbers"*** For example: $4 = 2+2$; $6 = 3+3$; $8 = 5+3$; $10 = 7+3 = 5+5$; etc. **To this day, it has not been proved** Goldbach's weak conjecture: ***"Every odd number bigger than 5 can be expressed as the sum of three prime numbers".*** **This conjecture was proved after about 270 years by Harald Helfgott, a Peruvian mathematician. Published in 2015.**

Appendix D4: Tasks in Discussion 4 of the 2018-intervention

<table>
<tr><th>Sub-Discussion</th><th>Task in Discussion 4 of the 2018-intervention</th></tr>
<tr><td>Discussion 4</td><td>Is it true that all universal statements have an infinite number of cases involved?
<table>
<tr><th>Statement 1</th><th>Statement 2</th></tr>
<tr><td>All divisions of natural numbers are exact divisions</td><td>All natural numbers smaller than 5 are smaller than 7</td></tr>
</table>
</td></tr>
</table>

Appendix D5: Tasks in Discussion 5 of the 2018-intervention

Sub-Discussion	Tasks in Discussion 5 of the 2018-intervention
Discussion 5	Is it true that all statements making reference to the **set of natural numbers** in its formulation will have an **infinite number of cases involved**? Let's see the following examples: 1) All natural numbers smaller than 5 are smaller than 7. 2) **All natural numbers such that added an even number of times together result in 6 are even numbers.** **¿Cuántos casos involucrados? ¿Infinitos?**
Task	**Classroom Episode 14 (see Appendix CE14)**

Appendix D6: Tasks in Discussion 6 of the 2018-intervention

<table>
<tr><th>Sub-Discussion</th><th>Tasks in Discussion 6 of the 2018-intervention</th></tr>
<tr><td>Discussion 6.1</td><td>In a piece of paper each teacher writes down an example of a true universal statement
(Hand over your example to EVV)
f) How do you know it is universal?
g) How do you know it is true?
(My two examples in the following slides)</td></tr>
<tr><td>Discussion 6.2.1:
My example 1</td><td>“0 is divisible by any natural number (different from cero)”
a) Is it a universal statement? How do you know?
b) Is it a true statement? Why?
c) What is it stated in this proposition?
d) What cases are involved in the statement? How many are they?
e) What characteristics does the justification have?</td></tr>
<tr><td>Discussion 6.2.2:
My example 2</td><td>“All natural numbers smaller than 5 are smaller than 7”
a) Is it a universal statement? How do you know?
b) Is it a true statement? Why?
c) What is it stated in this proposition?
d) What cases are involved in the statement? How many are they?
e) What characteristics does the justification have?</td></tr>
<tr><td>Discussion 6.2.3</td><td><table><tr><th>Statement 1</th><th>Statement 2</th></tr><tr><td>0 is divisible by any natural number (different from cero)</td><td>All natural numbers smalller than 5 are smaller than 7</td></tr></table>a) What similarities and differences can you find in my two statements?
b) If a statement is universal and true, what kind of Justification is expected in order to guarantee its truth value?</td></tr>
</table>

Appendix D7: Tasks in Discussion 7 of the 2018-intervention

<table>
<tr><th>Sub-Discussion</th><th>Tasks in Discussion 7 of the 2018-intervention</th></tr>
<tr><td>Discussion 7.1</td><td>In order to conclude the statement analysis
“All divisions of natural numbers are exact divisions”,
the teacher writes on the whiteboard:
Not all divisions of natural numbers are exact divisions
a) Is the second statement true or false?
b) How would you prove this truth value? Explain.
c) It is a universal statement?
d) In the statement is it stated that: no division of natural numbers is an exact division?
e) In the statement is it stated that: some divisions of natural numbers are exact?</td></tr>
<tr><td>Discussion 7.2</td><td>Statement 1: Not all divisions of natural numbers are exact divisions
Statement 2: It is not true that all divisions are exact divisions
Is there any difference between the two statements?
Do they state the same thing?
(They are equivalent statements)</td></tr>
<tr><td>Discussion 7.3</td><td>Statement 1: Not all divisions of natural numbers have a remainder 0
Statement 3: ……………………………………….
Write down a statement that states exactly the same to what’s stated in statement 1, but different to statement 2 seen in the previous slide
(Statement 3 equivalent to Statement 1)
Existential Statements</td></tr>
<tr><td>Discussion 7.4.1</td><td><table><tr><th>Statement 1</th><th>Statement 2</th></tr><tr><td>Some divisions are exact</td><td>Some divisions are not exact</td></tr></table>a) Is Statement 1 true? Why?
b) In Statement 1 is it stated that “Some divisions are not exact”?
c) The fact that Statement 1 is true implies that Statement 2 is true?</td></tr>
<tr><td>Discussion 7.4.2</td><td><table><tr><th>Statement 3</th><th>Statement 4</th></tr><tr><td>Some numbers divisible by 4 are even numbers</td><td>Some numbers divisible by 4 are not even numbers</td></tr></table>a) Is Statement 3 true? Why?
(to be justified. See next activity)
b) Statement 3 states that “Some numbers divisible by 4 are not even numbers”?
c) The fact that Statement 3 is true implies that Statement 4 is true?</td></tr>
<tr><td>Discussion 7.4.3</td><td><table><tr><th>Statement 1</th><th>Statement 2</th></tr><tr><td>Some divisions are exact</td><td>Some divisions are not exact</td></tr></table>a) Is Statement 1 true? Why?
b) In Statement 1 is it stated that “Some divisions are not exact”?
c) The fact that Statement 1 is true implies that Statement 2 is true?</td></tr>
<tr><td>Discussion 7.5</td><td>Not all natural numbers divisible by 2 are divisible by 3
a) Is it true or false?
b) Is it universal or existential?
c) What is exactly stated in this statement?
d) Write down two statements equivalent to the statement given
e) What characteristics does the justification that proves its truth value have?</td></tr>
</table>

Appendix D8: Tasks in Discussion 8 of the 2018-intervention

Sub-Discussion	Task in Discussion 8 of the 2018-intervention
Discussion 8	**Statement 1:** If A is divisible by B, then B is divisible by A. Is this statement true or false? Why? **Statement 2:** If A is divisible by B, then there exists cases in which B is divisible by A. Is this statement true or false? Why?

Appendix D9: Tasks in Discussion 9 of the 2018-intervention

<table>
<tr><th>Sub-Discussion</th><th>Tasks in Discussion 9 of the 2018-intervention</th></tr>
<tr><td>Discussion 9.1:
Activity 13</td><td>(see Appendix A13)
- Activity in PAIRs

Instructions:
1) Group the statements given in such a way that each group includes statements that state exactly the same thing.
2) Answer:
2.1) Are they universal or existential statements?
2.2) Are they true or false?</td></tr>
<tr><td>Discussion</td><td>Task:
Write down an example of a number that is divisible by itself. Explain why the number you suggested holds this condition.</td></tr>
<tr><td>Activity 14</td><td>Group activity[304]</td></tr>
<tr><td>Activity 15</td><td>Individual activity[305]</td></tr>
</table>

[304] See Appendix A14.
[305] See Appendix A15.

Appendix Co-A: Codes for the teachers' assumptions

Before Intervention	**During Intervention**	**After Intervention**
Example: bAG2 Assumption #2 Gessenia Assumption before intervention	Example: 1) dAA3[2] Assumption #2 Discussion 3 Andrea Assumption during intervention 2) dAL1[4] Assumption #4 Discussion 1 Lizbeth Assumption during intervention	Example: 1) aAGt3 Assumption #3 (her) Teaching Gessenia Assumption after intervention 2) aAAm2 Assumption #2 Meeting Andrea Assumption after intervention

Appendix AbI: The teachers' assumptions observed before the 2018-intervention

Code	Teacher	Assumptions observed before the intervention
bAA1	Andrea	*"All X are Y"* can be represented as set *X* included in set *Y*, with $Y - X \neq \emptyset$
bAA2	Andrea	If a UAS is (assumed to be) true, its converse is false
bAA3	Andrea	A UAS and its converse do not state the same
bAA4	Andrea	A counterexample is a conflicting example or a conflicting situation
bAA5	Andrea	Confirming examples are sufficient to prove UASs that involve infinite cases
bAA6	Andrea	*"There is X that is Y"* does not convey quantity
bAA7	Andrea	The true US *"All X are Y"* implies that *"There is X that is Y"* is true
bAA8	Andrea	*"All X are not Y" is the same as "Not all X are Y"*
bAG1	Gessenia	If a UAS is (assumed to be) true, its converse is true
bAG2	Gessenia	A UAS and its converse state the same
bAL1	Lizbeth	Confirming examples can *justify* that a UAS that involves infinite cases is true

Appendix AdI: The teachers' assumptions observed during the 2018-intervention

Code	Teacher	Assumptions observed during the intervention
dAA1[1]	Andrea	The true US *"All X are Y"* implies that *"Some X are Y"* is false
dAA1[2]	Andrea	Providing a "computation" is not sufficient justification to disprove a UAS
dAA1[3]	Andrea	Providing a counterexample can be considered a justification because the statement says ALL and if there is at least one that does not satisfy, then it is false
dAA1[4]	Andrea	*"Not all X are Y"* and *"It is false that all X are Y"* state the same
dAA1[5]	Andrea	A counterexample should "break" the statement
dAA1[6]	Andrea	I show my counterexample and prove that not all
dAA1[7]	Andrea	*"The X are Y"* is equivalent to *"All X are Y"*
dAA1[8]	Andrea	*"Some"* means *"from all, one group, but not all"*
dAA1[9]	Andrea	A counterexample must satisfy the first condition, but contradict the second condition of the statement
dAA1[9*]	Andrea	The counterexample must be always taken from the antecedent
dAA1[9**]	Andrea	The counterexample must satisfy the antecedent and not the consequent
dAA1[10]	Andrea	Confirming examples are insufficient to prove
dAA1[10]*	Andrea	Confirming examples are insufficient to prove a universal statement
dAA2[1]	Andrea	*"All X are Y"* can be represented as X included in Y and $Y - X$ may be empty
dAA2[2]	Andrea	If a UAS is (assumed to be) true, its converse *may* be true
dAA3[1]	Andrea	As long as the set of analysis is large, examples are not valid; if it is a small set, then they are
dAA3[2]	Andrea	*"For every x in N, P(x) is not Q"* is false if there is an *x* in *N* such that *P(x)* is *Q*
dAA7[1]	Andrea	*"Not all X are Y"* is equivalent to *"Some X are Y"*
dAA7[2]	Andrea	*"Not all X…"* is different from *"None X…"*
dAA7[3]	Andrea	*"Some X are Y"* necessarily implies that *"Some X are not Y"*
dAA7[4]	Andrea	One confirming example is sufficient to prove a "some-statement"
dAA7[5a]	Andrea	*"Some X are Y"* is false if it has a "counterexample"
dAA7[5b]	Andrea	*"Some X are not Y"* is false if it has a "counterexample"
dAA7[6]	Andrea	*"Some X are not Y"* is false if it is impossible to find an example of *X* that is not *Y*
dAA7[7]	Andrea	"Some" means "at least one, and maybe all"
dAA7[8]	Andrea	If *"Some X are Y"* and *"All X are Y"* are both true, then *"Some X are not Y"* is false
dAA7[9]	Andrea	If *"Some X are Y"* is true and *"All X are Y"* is false, then *"Some X are not Y"* is true
dAA7[10]	Andrea	The negation of *"All X are Y" ("All X are not Y")* is *"Some X are not Y" ("Some X are Y")* (her DSS-approach)
dAA7[11]	Andrea	*"All X are not Y"* is true if there is an *X* that is not *Y*
dAA7[12]	Andrea	*"No X is Y"* and *"No X is not Y"* refers to everything that is not *X*
dAA7[13]	Andrea	*"No X is Y"* and is represented as two disjoint sets *X* and *Y*
dAA7[14]	Andrea	*"No X is Y"* is equivalent to *"All X are not Y"*
dAA7[15]	Andrea	The negation of a *"some-statement"* is an *"all-statement",* and never a *"no-statement"*
dAA7[16]	Andrea	The negation of *"Some X are Y"* (*"Some X are not Y"*) is *"All X are not Y"* (*"All X are Y"*)
dAA7[17]	Andrea	The negation of "no…" is "some…"
dAA7[18]	Andrea	The negation of *"No X is Y"* is *"Some X are not Y"* (her DSS-approach)
dAA7[19]	Andrea	The negation of *"All X are not Y"* is *"Some X are Y"*
dAA7[20]	Andrea	*"No X is Y"* is different from *"All X are not Y"*
dAA7[21]	Andrea	The negation of *"Some X are Y"* is *"No X is Y"*
dAA7[22]	Andrea	The negation of *"No X is Y"* is *"Some X are Y"*
dAA7[23]	Andrea	*"No X is not Y"* is equivalent to *"All X are Y"*
dAA7[24]	Andrea	The negation of *"No X is not Y"* = The negation of *"All X are Y"* = *"Some X are not Y"*
dAA7[26]	Andrea	*"No X is not Y"* is represented as two disjoint sets *X* and *Y*

dAA9[1]	Andrea	*"All X are not Y"* is false if there is an example of *X* that is *Y* (a counterexample)
dAA9[2]	Andrea	*"There does not exist X that is Y"* is false if there is an *X* that is *Y* (i.e., because *"There exist X that is Y"* is true)
dAA9[3]	Andrea	There does not exist *X* = No *X*
dAA9[4]	Andrea	*"There does not exist X that is Y"* is the same as *"No X is Y"*
dAA9[5]	Andrea	*"There does not exist X that is not Y*" is false if there exist an *X* that is *Y.*
dAA9[6]	Andrea	*"There does not exist X that is not Y" is the same as "No X is not Y"*
dAA9[7]	Andrea	*"There does not exist X that is not Y" is false if there is an X that is not Y*
dAA9[8]	Andrea	*"There does not exist X that is not Y" is the same as "No X is Y"*
dAA9[9]	Andrea	When the statement is universal and true with infinite cases involved, an example is not a valid justification, unless it is a generic example
dAG1[1]	Gessenia	Providing a "computation" is not sufficient justification to disprove a UAS
dAG1[2]	Gessenia	The statement *"All X are Y"* is false because *not all X are Y*
dAG1[3]	Gessenia	A counterexample should "break" the statement
dAG1[4]	Gessenia	Counterexamples should not satisfy at least one condition in the statement
dAG1[5]	Gessenia	Counterexamples should contradict the statement
dAG2[1]	Gessenia	St49 (a specific UAS) and its converse do not state the same
dAG2[2]	Gessenia	In the UAS *"All X are Y"*, *X* and *Y* have the same elements (*X* = *Y*)
dAG2[3]	Gessenia	In the UAS *"All X are Y"*, *X* and *Y* may not have the same elements
dAL1[1]	Lizbeth	A counterexample "breaks" the statement
dAL1[2]	Lizbeth	It (a confirming example) cannot be a counterexample because it supports St1 (the statement)
dAL1[3]	Lizbeth	Confirming examples *prove* a UAS with infinite cases involved, though they do not *guarantee* that the UAS is true
dAL1[4]	Lizbeth	It (an irrelevant example) cannot be a counterexample because it does not satisfy the first condition of St35 (the statement)
dAL1[5]	Lizbeth	The order in the statement is important
dAL3[1]	Lizbeth	Verifying some examples does not *guarantee* that a UAS is true. All examples need to be verified in order to do that
dAL3[2]	Lizbeth	In order to *justify* that a statement is true, we should check each case involved in the statement
dAL3[3]	Lizbeth	If we do not find any counterexamples for the US-conjecture, then it is a mathematical truth

Appendix AaI: The teachers' assumptions observed after the 2018-intervention

Code	Teacher	Assumptions observed after the intervention
aAAt1	Andrea	*"Some X are Y"* is false because it is impossible to find an example of *X* that is *Y*
aAAt2	Andrea	When I say "all" or "none" and I give a counterexample to prove that it is false, that is valid
aAAm1	Andrea	A verbalized semi-general counterexample is valid to disprove a US
aAAm2	Andrea	If a UAS is (assumed to be) true, its converse *may* be false
aAAm3	Andrea	Given two disjoint sets *X* and *Y*, *Y* represents the set of the elements that are "not *X*"
aAGt1	Gessenia	I need to see; otherwise, it is not evident. Show me a counterexample
aAGt2	Gessenia	A counterexample should verify the first condition, but contradict the second condition of St45
aAGt3	Gessenia	When it says ALL and it is not true, it is sufficient to show a counterexample. There might be many counterexamples, but one is sufficient to disprove an "ALL-statement"
aAGm1	Gessenia	A true UAS does not necessarily imply that its converse is true
aAGm2	Gessenia	A UAS does not assert its converse, but it does not deny it either

Appendix Sts: Statements seen before, during and after the 2018-intervention

Statements seen *before* the intervention

Where	Code	Statement	Type	Specific type
EI1 Written test (task II, T1)	St1	*The sum of two even numbers is an even number*	UAS	For all x, y, P(x;y)
EI1 Written test (task II, T2)	St2	*All «Vallejo» numbers are even numbers*	UAS	ALL-S
	St3	*Some «Vallejo» numbers are even numbers*	EAS	SOME-S
	St4	*All even numbers are «Vallejo» numbers*	UAS	ALL-S
	St5	*If a number is «Vallejo», then it is an even number*	UCS	IF-THEN-S
	St6	*No «Vallejo» number is an even number*	UNS	NO-S
	St7	*The even numbers are not «Vallejo» numbers*	UNS	THE-S
	St8	*Some «Vallejo» numbers are not even numbers*	ENS	SOME-S
	St9	*Every «Vallejo» number is an even number*	UAS	EVERY-S
	St10	*The «Vallejo» numbers are even numbers*	UAS	THE-S
	St11	*If a number is even, then it is a «Vallejo» number*	UCS	IF-THEN-S
	St12	*There are «Vallejo» numbers that are even numbers*	EAS	THERE-ARE-S
	St13	*There are no «Vallejo» numbers that are not even*	Negation of ENS	Simple implicit negation of THERE-ARE-S
	St14	*Not all «Vallejo» numbers are even numbers*	Negation of UAS	Simple implicit negation of ALL-S
EI1 Written test (task II, T3)	St15	*If we multiply any natural number by 5, we get a number with its unit-digit equal to 0 or 5*	UCS	IF-THEN-S

Statements seen *during the first part* of the intervention

Where	Code	Statement	Type	Specific type
Activity 6 (slide 50)	T1	*In a division (FWM distribution) the remainder must be smaller than the number of people to whom the objects will be distributed, considering zero as the smallest remainder and the number prior to the number of people as the highest remainder.*	UAS	
Discussion (Slide 60)	T2	*In a division where the number of objects to distribute is smaller than the number of people to whom the objects will be distributed, the result (quotient) will be zero and the remainder will be the number of objects.*	UAS	
Discussion (slide 63)	T3	*A number is not divisible by a bigger number, considering that the first number is bigger than 0.*	UNS	
Discussion (slide 63)	T4	*Zero is divisible by any natural number bigger than zero.*	UAS	ANY-S
Discussion (slide 63)	T5	*All numbers are divisible by 1*	UAS	ALL-S
Discussion (slide 63)	T6	*All numbers are divisible by themselves, except for zero.*	UAS	ALL-S

Statements seen *during* Discussion 1

Where	Code	Statement	Type	Specific type
Intro D1 (A11)	St16	*All divisions of natural numbers are exact divisions*	UAS	ALL-S
	St17	*0 is divisible by any natural number (different from 0)*	UAS	FOR-ALL-S
	St18	*All natural numbers smaller than 5 are smaller than 7*	UAS	ALL-S
	St19	*It is not true that all divisions of natural numbers have a remainder equal to zero*	Negation of UAS	Implicit negation of ALL-S
	St20	*All numbers divisible by 4 are divisible by 2*	UAS	ALL-S
	St21	*Not all numbers divisible by 4 are divisible by 2*	Negation of UAS	Simple implicit negation of ALL-S
	St22	*All numbers divisible by 2 are not divisible by 4*	UNS	ALL-S
	St23	*No natural number is divisible by 4 and 6*	UNS	NO-S
	St24	*Some divisions are exact*	EAS	SOME-S
	St25	*Some divisions are not exact*	ENS	SOME-S
	St26	*Some numbers divisible by 4 are even numbers*	EAS	SOME-S
	St27	*Some numbers divisible by 4 are not even numbers*	ENS	SOME-S
	St28	*There do not exist natural numbers divisible by 2 that are divisible by 3*	Negation of EAS	Simple implicit negation of THERE-EXIST-S
	St29	*If a whole and fair distribution is also maximal, then the remainder is zero*	UCS	IF-THEN-S
	St30	*If a distribution is FWM, then the number of objects left is smaller than the number of people*	UCS	IF-THEN-S
	St31	*If a number is divisible by 2, then it is divisible by 4*	UCS	IF-THEN-S
	St32	*If a number is divisible by 4, then it is not divisible by 2*	UCS	IF-THEN-S
	St33	*Some numbers divisible by 4 are divisible by 8*	EAS	SOME-S
D1.0	St16		UAS	ALL-S
D1.1	St16		UAS	ALL-S
	St34	*Not all divisions of natural numbers are exact divisions*	Negation of UAS	Simple implicit negation of ALL-S
D1.2	St16		UAS	ALL-S
D1.3	St16		UAS	ALL-S
D1.4.1	St16		UAS	ALL-S
D1.4.2	St16		UAS	ALL-S
D1.5	St35	*The people present in this room are minors*	UAS	THE-S
	St36	*Some people in this room are minors*	EAS	SOME-S
D1.6.1	St35		UAS	THE-S
D1.6.2	St16		UAS	ALL-S
D1.7	St37	*All numbers divisible by 6 are divisible by 3*	UAS	ALL-S
After D1.7	St38*	*All chickens are animals*	UAS	ALL-S
	St39*	*All dogs are dinosaurs*	UAS	ALL-S
	St2*		UAS	ALL-S
Recap D1 (includes CE10)	St40*	*All even numbers end in an odd digit*	UAS	ALL-S
	St41*	*All numbers that end in digit 3 are divisible by 4*	UAS	ALL-S
	St42*	*All numbers divisible by 3 are even numbers*	UAS	ALL-S
	St43	*All divisions of natural numbers have a remainder zero*	UAS	ALL-S
	St44	*The whole and fair distributions must have a remainder zero so that they are also maximal*	UCS	Implicit IF-THEN-S

Statements seen *during* Discussion 2

Where	Code	Statement	Type	Specific type
D2.1	St45	*If a distribution is FWM, then the remainder is zero*	UCS	IF-THEN-S
	St46	*Some FWM distributions have a remainder zero*	EAS	SOME-S
D2.2	St30		UCS	IF-THEN-S
D2.3	St37		UAS	ALL-S
	St47	*If a number is divisible by 6, then it is divisible by 3*		
D2.4 (CE13.2)	St45		UCS	IF-THEN-S
D2.4.1	St45		UCS	IF-THEN-S
	St48	*If the remainder is zero, then the distribution is FWM*	UCS	IF-THEN-S
	St49*	*If it is a person, then it is a mortal*	UCS	IF-THEN-S
	St50*	*If it is a mortal, then it is a person*	UCS	IF-THEN-S
D2.4.2	St51	*If a number is even, then it is a FAST number*	UCS	IF-THEN-S
	St52	*If a number is FAST, then it is even*	UCS	IF-THEN-S
	St53	*If it is a human being, then it is mortal*	UCS	IF-THEN-S
	St54	*If it is a mortal, then it is a human being*	UCS	IF-THEN-S
	St55	*If a person is from Lima, then s/he is Peruvian*	UCS	IF-THEN-S
	St56	*If a person is Peruvian, then s/he is from Lima*	UCS	IF-THEN-S
	St57*	*If it is a mammal, then it nurses*	UCS	IF-THEN-S
D2.4.3	St58	*If the unit-digit of a number is 0, then the number is divisible by 10*	UCS	IF-THEN-S
	St59	*If a number is divisible by 10, then the unit-digit of the number is 0*	UCS	IF-THEN-S
	St60*	*If s/he is Peruvian, then s/he was born in Peru*	UCS	IF-THEN-S
	St61*	*If a number end in an even digit, then it halves*	UCS	IF-THEN-S
	St62*	*If a number end in 5 or 0, then it is divisible by 5*	UCS	IF-THEN-S
	St63*	*If a number is divisible by 5, then it ends in 5 or 0*	UCS	IF-THEN-S
D2.4.4	St64	*If the unit-digit of a number is 0, then the number is divisible by 5*	UCS	IF-THEN-S
	St65	*If a number is divisible by 5, then its unit-digit is 0*	UCS	IF-THEN-S
JV3	St66	*If a distribution is whole and fair, then it is FWM zero*	UCS	IF-THEN-S
	St67	*If a distribution is FWM, then it is fair*	UCS	IF-THEN-S
	St68	*If a distribution is FWM, then it is whole*	UCS	IF-THEN-S
	St69	*If a whole and fair distribution has a remainder zero, then it is FWM*	UCS	IF-THEN-S

Statements seen *during* Discussion 3

Where	Code	Statement	Type	Specific type
Activity 12 (A12)	St70	*Every time we work with a bigger number, we will always get more true sentences than in the previous cases*	UAS	FOR-EVERY-S
D3.1	St71	*When we divide a number C by a number A bigger than C, the result is zero and the remainder is equal to C*	UAS	WHEN-THEN-S
D3.2	St72	*All palindrome numbers are divisible by 11*	UAS	ALL-S
Recap	St73	*All men are liars*	UAS	ALL-S
	St74*	*All one-digit numbers divisible by 6 are divisible by 3*	UAS	ALL-S
Task	St75	*For all* n *natural number (different from zero), the following expression:* $$1 + 1141n^2$$ *never produces a perfect square*	UNS	FOR-ALL-S
D3.3	St76	*All natural numbers of the form* $2^{2^n} + 1$, *with n a natural number, are prime numbers*	UAS	ALL-S
D3.4	St77	*All even numbers bigger than 2 can be written as the sum of two prime numbers*	UAS	ALL-S
	St78	*All odd numbers bigger than 5 can be expressed as the sum of three prime numbers*	UAS	ALL-S

Statements seen *during* Discussion 4

Where	Code	Statement	Type	Specific type
D4	St74		UAS	ALL-S
	St16		UAS	ALL-S
	St79	*All natural numbers smaller than 5 are smaller than 7*	UAS	ALL-S

Statements seen *during* Discussion 5

Where	Code	Statements	Type	Specific type
D5	St74		UAS	ALL-S
	St75		UNS	FOR-ALL-S
	St79		UAS	ALL-S
	St80	*All natural numbers such that added an even number of times give 6 are even*	UAS	ALL-S
CE14	St81	*All numbers divisible by 4 are divisible by 8*	UAS	ALL-S

Statements seen *during* Discussion 6

Where	Code	Statement	Type	Specific type
D6.1	St82*	*All even numbers are divisible by 2*	UAS	ALL-S
	St83*	*All even numbers are multiples of 2*	UAS	ALL-S
	St84*	*All natural numbers that end in an even digit are infinite*	UAS	ALL-S
D6.2.1	St17		UAS	FOR-ALL-S
D6.2.2	St79		UAS	ALL-S
D6.2.3	St17		UAS	FOR-ALL-S
	St79		UAS	ALL-S

Statements seen *during* Discussion 7

Where	Code	Statement	Type	Specific type
D7.1	St16		UAS	
	St34		Negation of UAS	Simple implicit negation of ALL-S
	St24	*Some divisions of natural numbers are exact divisions*	EAS	SOME-S
	St85	*No division of natural numbers is an exact division*	UNS	NO-S
	St25	*Some divisions of natural numbers are not exact divisions*	ENS	SOME-S
After D7.1	St86*	*All natural numbers that end in digit 0 are multiples of 5*	UAS	ALL-S
	St87	*Some natural numbers that end in digit 0 are not multiples of 5*	ENS	SOME-S
D7.2	St34		Negation of UAS	Simple implicit negation of ALL-S
	St88	*It is not true that all divisions of natural numbers are exact divisions*	Negation of UAS	Implicit negation of ALL-S
D7.3	St89	*Not all divisions of natural numbers have a remainder 0*	Negation of UAS	Simple implicit negation of ALL-S
	St90*	*Some divisions of natural numbers do not have remainder 0*	ENS	SOME-S
D7.4.1	St24		EAS	SOME-S
	St25		ENS	SOME-S
	St86		UAS	ALL-S
	St91	*Some natural numbers that end in digit 0 are multiples of 5*	EAS	SOME-S
D7.4.2	St92	*All numbers divisible by 4 are even*	UAS	ALL-S
	St26		EAS	SOME-S
	St27		ENS	SOME-S
	St93*	*All got an "A" in the test*	UAS	ALL-S
	St94*	*Some got an "A" in the test*	EAS	SOME-S
	St95*	*Some did not get an "A" in the test*	ENS	SOME-S
	St96*	*All teachers in the classroom are women*	UAS	EVERY-S
	St97*	*Some teachers in the classroom are women*	EAS	SOME-S
	St16		UAS	ALL-S
	St86		UAS	ALL-S
	St87		EAS	SOME-S
	St98*	*Some numbers that end in digit 3 are multiples of 3*	EAS	SOME-S
	St99*	*All numbers that end in digit 3 are multiples of 3*	UAS	ALL-S
D7.4.3	St24		EAS	SOME-S

	St25		ENS	SOME-S
	St100*	*Some even numbers are «Vallejo» numbers*	EAS	SOME-S
	St101*	*Some even numbers are not «Vallejo» numbers*	ENS	SOME-S
	St102*	*If a number is not even, then it is odd*	UCS	IF-THEN-S
D7.5	St103	*Not all natural numbers divisible by 2 are divisible by 3*	Negation of UAS	Simple implicit negation of ALL-S
After D7.5	St104*	*All natural numbers divisible by 2 are not divisible by 3*	UNS	ALL-S
	St105*	*Some natural numbers divisible by 2 are divisible by 3*	EAS	SOME-S
	St106	*Some natural numbers divisible by 2 are not divisible by 3*	ENS	SOME-S
	St107	*All natural numbers divisible by 2 are divisible by 3*	UAS	ALL-S
	St108	*No odd number is an Innova number*	UNS	NO-S
	St109	*All odd numbers are not Innova numbers*	UNS	ALL-S
	St110*	*All students are inside the classroom*	UAS	ALL-S
	St111*	*No student is outside the classroom*	UNS	NO-S
	St112	*No student is not inside the classroom*	UAS	Negative NO-S
	St113	*Some odd numbers are not Innova numbers*	ENS	SOME-S
	St114	*Some odd numbers are Innova numbers*	EAS	SOME-S
	St115	*No odd number is not an Innova number*	UAS	Negative NO-S
	St116	*All odd numbers are Innova numbers*	UAS	ALL-S
	St117	*No man is not a living being*	UAS	Negative NO-S
	St118	*All men are living beings*	UAS	ALL-S

Statements seen *during* Discussion 8

Where	Code	Statement	Type	Specific type
D8	St119	*If A is divisible by B, then B is divisible by A*	UCS	IF-THEN-S
	St120	*If A is divisible by B, then there exist cases where B is divisible by A*		IF-THEN-S

Statements seen *during* Discussion 9

Where	Code	Statement	Type	Specific type
D9 (A13)	St89		Negation of UAS	Simple implicit negation of ALL-S
	St121	*All divisions of natural numbers do not have a remainder zero*	UNS	ALL-S
	St43		UAS	ALL-S
	St122	*There exist divisions of natural numbers that have a remainder zero*	EAS	THERE-EXIST-S
	St123	*There do not exist divisions of natural numbers that have a remainder zero*	Negation of EAS	Simple implicit negation of THERE-EXIST-S
	St124	*There exist divisions of natural numbers that do not have a remainder zero*	ENS	THERE-EXIST-S
	St125	*No division of natural numbers has a remainder zero*	UNS	NO-S
	St126	*There do not exist divisions of natural numbers that do not have a remainder zero*	Negation of ENS	Simple implicit negation of THERE-EXIST-S
	St127	*No division of natural numbers does not have a remainder equal to zero*	UAS	Negative NO-S
Activity	St128	*All natural numbers (except for zero) are divisible by themselves*	UAS	ALL-S
(A14)	St129	*All natural numbers are divisible by 1*	UAS	ALL-S
(A15)	St130	*The numbers divisible by 4 are infinite*	UAS	THE-S

Statements seen *after* the intervention

Where	Code	Statement	Type	Specific type
M2	St129		UAS	ALL-S
IS3	P1	*In FWMD with #obj < #people, each person gets 0 and r = #obj*	UAS	
IS4	P2	*#obj left < #people, for all FWMD*	UAS	FOR-ALL-S
M4; TS7	St45	*If a distribution is fair, whole and maximum, then there will be 0 objects left*	UCS	IF-THEN-S
	St131	*Some divisions by 4 have a remainder equal to 7*	EAS	SOME-S
M5	St25		ENS	SOME-S
	St132	*All natural numbers smaller than 3 are smaller than 6*	UAS	ALL-S
	St133	*If a distribution is fair, then it is FWM*	UCS	IF-THEN-S
M6	St134	*There are no numbers that are not bigger than 6*	Negation of ENS	Simple implicit negation of THERE-ARE-S
	St135	*The divisions of the first five natural numbers by 3 are exact divisions*	UAS	THE-S
	St136	*No number is odd*	UNS	NO-S
A[TS6]	P3	*Max #obj left = #people – 1, for all FWMD*	UAS	
M7	St135		UAS	THE-S

	St137	*Some numbers when divided by 3 result in exact divisions*	EAS	SOME-S
M9	St45			
	St138	*No number greater than 7 is a palindrome number*	UNS	NO-S
M10	St139	*All natural numbers bigger than 5 are RAINBOW numbers*	UAS	ALL-S
M11	St140	*All Peruvians are BLUMEN*	UAS	ALL-S
	St141	*Some RAINBOW numbers are smaller than 3*	EAS	SOME-S
G[TS10]; A[TS10]	St142	*Some natural numbers are divisible by 4*	EAS	SOME-S
	St143	*All natural numbers are divisible by 4*	UAS	ALL-S
M12	St144	*No even number is a VOILA number*	UNS	NO-S
	St145	*No even number is not a VOILA number* (UAS)	UAS	Negative NO-S
	St146	*There do not exist even numbers that are VOILA*	Negation of EAS	Simple implicit negation of THERE-EXIST-S
	St147	*There do not exist natural numbers that cannot be divided by 2*	Negation of ENS	Simple implicit negation of THERE-EXIST-S
L[TS10]; G[TS11]	P4	*All numbers are divisible by themselves, except for 0*	UAS	ALL-S
M13	St128			
A[TS11]	P5			
TS12	St70		UAS	FOR-EVERY-S
	St148	*The numbers divisible by 3 are infinite*		
M14	St37		UAS	
TS13	St149	*All numbers divisible by 3 are bigger than 3*	UAS	ALL-S
TS14	St150	*Some natural numbers are not divisible by 4*	ENS	SOME-S
	St151	*All natural numbers divisible by 4 are divisible by 8*	UAS	ALL-S
	St152	*No natural number is divisible by 4*	UNS	NO-S
	St153	*(Consider the numbers 0, 3 and 9)* *All these numbers are divisible by 1*	UAS	ALL-S
	St154	*Some numbers divisible by 4 are divisible by 8*	EAS	SOME-S
M17		*Introductory Test Review*		