



Titel/Title: When Do Numbers Count? The Differential Impact of the PISA Rating and Ranking on Education Policy in Germany and the US

Autor*innen/Author(s): Kerstin Martens and Dennis Niemann

Veröffentlichungsversion/Published version: Postprint

Zeitschriftenartikel/Journal article

Empfohlene Zitierung/Recommended citation:

Kerstin Martens & Dennis Niemann (2013) When Do Numbers Count? The Differential Impact of the PISA Rating and Ranking on Education Policy in Germany and the US, *German Politics*, 22:3, 314-332, DOI: 10.1080/09644008.2013.794455

Verfügbar unter/Available at:

(wenn vorhanden, bitte den DOI angeben/please provide the DOI if available)

<https://doi.org/10.1080/09644008.2013.794455>

Zusätzliche Informationen/Additional information:

The author can be contacted at: kerstin.martens@sfb597.uni-bremen.de

When Do Numbers Count? The Differential Impact of the PISA Rating and Ranking on Education Policy in Germany and the US

KERSTIN MARTENS and DENNIS NIEMANN

Why do international comparisons have an impact on some countries while other countries do not respond? This article examines the power of international ratings and rankings (R&R) using the OECD's PISA study and its differential impact on education policymaking as a case study. It argues that international R&R have an impact when two conditions are simultaneously fulfilled: the evaluated topic is framed as crucial in the national discourse, and a substantial gap between national self-perception and the empirical results can be observed. After assessing the media impact of PISA on 21 OECD countries, the theoretical argument is illustrated by an evaluation of the differing reactions of two similarly poorly performing countries: Germany and the US. While the German system of secondary education was strongly affected by the international comparison, beginning with the first PISA study in 2000, and underwent comprehensive changes, the US first responded noticeably in the public and the political discourse only in 2010 to its below-average ranking.

INTRODUCTION

Comparative depictions have become prominent phenomena in modern globalised societies. In particular, evaluative illustrations in the form of ratings and rankings (R&R) increasingly occur today in many policy fields and issue areas. Such R&R are also applied to states, state institutions and state performance in specific policy fields.¹ Often, the outcomes of such R&R influence actors' future decisions, be they of governmental representatives, market actors or individuals. Comparisons are popular in modern societies because they provide succinct information quickly and in an easily understandable way: without being an expert in the specific field anybody can understand that being ranked #5 is different from #77, or that a positive value in any rating differs from a negative. R&R are not a novelty. However, what is fairly new is that results of R&R initiate public and political debates because they create an air of competition around state performance or policies by attributing relative positions. Although R&R usually do not provide direct recommendations or policy proposals, they can at least draw attention to possible solution strategies or best practice examples in a specific problem field and, hence, open a window of opportunity for policy reforms.

A particularly prominent R&R in the field of education is the OECD's Programme for International Student Assessment (PISA). When the initial results from this cross-

national comparison of the knowledge and skills of secondary school students were released in early December 2001, a contentious and continuing debate about education was unleashed in Germany: while the country had long prided itself on its education system with its contributions to Western science and philosophy, the international comparative data empirically revealed that the expected superiority of the German education system appeared to be no more than mere mediocrity. In essence, over ten years of experience with PISA unleashed a comprehensive reform initiative in the German secondary education system, which is still ongoing today.

Unlike in Germany, PISA remained almost unnoticed and triggered little public discussion or reform in education policy in the US – despite the fact that Americans are faced with similarly poor results as Germany. Proud of being the leading nation in tertiary education, the prime destination for international students and with eight universities among the top ten of the Shanghai Ranking,² the US public and a large part of policymakers do not seem to be concerned when it comes to the evaluation of its secondary education system through PISA. Only with the publication of the fourth and latest PISA results in 2010 did the US finally become aware of this comparative study. It was in particular, the discrepancy between the Chinese in Shanghai, tested for the first time in this PISA study, and the US students which provided – as education minister Arne Duncan put it – the ‘wake-up call’ to improve the US school system.³ In fact, the first three PISA rounds already indicated the flaws of its secondary education, but the US apparently preferred to push the ‘snooze button’ when the OECD’s alarm clock rang.

Although the PISA study is not the first international comparative education study, it is by now the largest internationally executed study. Surveying the competencies and skills of 15-year-olds in reading, mathematics and science near the end of compulsory schooling, PISA is now administered in 65 countries.⁴ It first began in 2000 and has been conducted every three years since with the results being published the following year. Developed and carried out by the OECD, PISA is obviously not ideologically unbiased but rather evaluates education from an economic perspective and promotes, according to this paradigm, related learning techniques.⁵ PISA focuses primarily on the output of education systems and seeks to assess how well students are able to apply knowledge and skills learned in school to their future working life and thus contribute to the prosperity of the national economy.⁶ In addition, PISA has also made the OECD an ‘eminence grise’ in education policy since today the International Organisation (IO) is regarded as an authority in education.⁷

The question we are dealing with in this article is: under what circumstances can international R&R, such as PISA, have an influence on national reactions in terms of public response that may lead to substantial reforms in the education sector? We propose that two conditions must be fulfilled simultaneously: a substantial gap between national self-perception and the empirical results (‘perception’) can be observed, *and* the evaluated topic is framed as crucial for state purposes (‘framing’). With our case, we thus contribute to understanding under what conditions international evaluative assessments are influential processes that impact on national settings. By examining the particularly prominent case of the OECD’s PISA study as an example of R&R, this article analyses the governance by comparison that IOs are able to demonstrate through direct quantitative comparisons. In this way, our study

also contributes to the ongoing research about the ability of IOs to influence national policymaking.⁸

Our study presents one of the few proposals for understanding the differential impact R&R may have. Although these comparative depictions enjoy continuous application, surprisingly few academic studies deal with analysing the reach and impact of R&R. While some contributions investigate more generally the prestige and power of quantification in the modern world, the political impact that R&R may have is rarely conceptualised.⁹ Although R&R are popular in the field of (higher) education, it was mainly PISA that triggered works on how and by what means these comparative depictions create attention and competition amongst participating members, particularly with regard to German underachievement in the OECD study.¹⁰

In the next section of this article, we develop a heuristic model for understanding the power of R&R in different national contexts highlighting *perception* and *framing* as necessary conditions for international R&R to be decisive at the national level. We do not seek to establish causal linkages nor do we claim to present a dense theoretical model. Rather, we suggest conditions that, in our view, account for the divergent responses to R&R. Following that, the diversity of national reactions to PISA is exemplified by the analysis of media coverage in 21 contributing OECD countries. The aim of this section is to show that the extent of public reactions related to PISA cannot be linked directly to the actual position of a country in the PISA ranking. Even countries with a similar performance vary substantially in their public reactions.

By examining two cases of poorly performing PISA countries that vary in their public and political responses – Germany and the US – the different reaction schemes are illustrated and linked to the theoretical assumptions in the third section of this article. In the German case, we show that the PISA study has managed to transform poor results into activation of broad secondary education reform dynamics by revealing a gap between self-perception and evidence as well as by generating a link to other crucial issues of state performance. In contrast, these mechanisms have failed in the case of the US with regard to PISA. The international comparative study did not provide new evidence on the relatively poor secondary school system and could not trigger related nationwide reforms. Only recently, with the fourth PISA study published in late 2010, did the media and policymakers in the US begin to problematise the importance of its poor results for its economic future.

Our article examines the two-level game of exploring the impact of international assessments on national settings. By focusing on how the OECD's PISA study influences policymaking in the field of education policy, we deal with hard cases.¹¹ The OECD as an IO does not comprise any binding instruments but relies on soft means such as its expert status when addressing states with its reports;¹² the field of education is generally believed to be anchored firmly in the national context and thus more 'resistant' to international influences than other policy fields. Additionally, secondary education policy has no need to be coordinated at the international level since usually no dilemma situation, as in other policy fields, prevails.

Our study presents findings of comparative work on two country cases following a most-similar-systems design. In both states secondary education is decentralised.¹³ It is the German Land and the US state that are formally in charge of education as a policy field; thus the governments in Berlin and Washington DC cannot impose nationwide

reform processes. Nevertheless, the need for national coordination in education as a reaction to poor PISA results should be equally apparent. In addition, Germany and the US are both founding members of the OECD, thus their experience in dealing with the initiatives of this IO in the form of recommendations to and evaluations of national policies can be assumed to be similar. They differ, however, in the reception of their individual PISA results. We analyse their reactions to the international R&R by comparing the pre-PISA (– 1999) and post-PISA (2000 –) era.

GOVERNANCE BY COMPARISON – WHEN DO R&R HAVE AN INFLUENCE?

How can differing reactions to the OECD's PISA study be explained? Or to put it more generally: when do R&R lead to strong public responses and subsequent reforms? We set up a parsimonious heuristic framework for assessing the impact of R&R in national policymaking processes. We argue that there are basically two approaches in which R&R have an impact in the national context: if they reveal a gap between national self-perception and empirical evidence that is too substantial to be ignored (*perception*), and – at the same time – if the actors who have installed the R&R manage to frame the topic under evaluation as crucial for state purposes (*framing*). Only the complementary use of both, perception and framing, in our theoretical framework allows for the comprehensive assessment of national reactions toward international initiatives. This framework combines the governance perspective with agenda-setting theory.

From a *governance* viewpoint, comparisons illustrate that power is not wielded solely by traditional top-down regulatory activities but also by apparently soft comparative R&R such as those of IOs.¹⁴ The parties evaluated are implicitly pressured to converge towards those practices, forms of organisation or behaviour that are best evaluated in line with the specific criteria in the respective framework of comparison. In this regard, R&R can provide a roadmap of what has to be done. These soft forms of governance by comparison emerging in the transnational sphere have the potential to influence established institutionalised procedures at the national level by creating standards, which then produce pressure for improvement. R&R are, therefore, representative of the debated shift from government to governance.¹⁵

According to *agenda-setting theory* at the national level, comparisons carried out by IOs account for the international circulation of standards, ideas and solutions for problems.¹⁶ In this respect, IOs become part of the so-called policy stream since the experts working in them create the indicators for R&R and hence define the benchmarks against which national policies are evaluated. Through R&R, IOs develop and promote policy ideas that float around and transcend national borders. Based on policies promoted through IOs and translated into R&R, the national discourse on a topic can be influenced by the narrowing down of the set of subjects that actually becomes the focus of attention.¹⁷ In this model, statistical indicators are seen as one way of socially constructing an economic or social problem. They might reveal something that was not regarded as problematic before. In the case where the IO reaches policymakers and the public at the same time, R&R may open up a policy window in the problem stream. Agenda-setting theory illustrates the interaction between the

media, public opinion and policymakers showing why and how public issues rise and fall in importance over time.¹⁸

The Substantial Gap between Self-perception and Results

From a rationalist perspective, international comparative R&R can be seen to reduce transaction and information costs in liberal markets and societies. This also holds true for education policy: because of increasing demands on labour forces operating within a global market, states are constantly in need of improving their education systems.¹⁹ In this respect, comparative studies are an instrument for revealing the 'best practices' available that then can be emulated by poorer performers. From the point of view of states, IOs such as the OECD are considered suitable bodies for the task of evaluating education systems through a comparative perspective: they provide the capacities needed for developing and applying objective criteria for an international assessment of different states and act 'with a degree of autonomy'.²⁰

To be evaluated does not necessarily mean that self-perception is simultaneously created by the specific ranking; R&R themselves do not constitute a whole new perception on an issue. Rather, pre-existing self-impressions regarding expectations about the performance collide with empirical results of R&R. If there is a divergence between the expected results and empirical findings, R&R might lead to a re-evaluation of the self-impression regarding the evaluated topic or trigger reform pressure in order to meet future expectations. With the PISA ranking, the OECD manages to easily quantify and reveal the possible gap between self-impression and empirical evidence.

R&R are most likely to have an impact at the national level if they uncover a substantial negative gap between the pre-existing self-perception and the actual empirical results. In this context, a substantial negative gap can be defined as a situation in which the actual performance deviates from the expectation in a sense that the outcome does not maintain the assigned task in terms of either effectiveness or efficiency. R&R can serve as a focal point, drawing attention to a new problem.²¹ Hence, comparative evaluations conducted by IOs can show if and to what extent a country is lagging behind current international standards. Furthermore, R&R provide points of orientation regarding what measures for improvement have to be taken into account in order to overcome deficits. In the case of the OECD and PISA, the IO has to disclose a mismatch between performance and national self-perception. PISA establishes benchmarks for secondary education systems that require comprehensive reforms to meet the standards of other industrialised states.

Besides revealing a gap and defining a problem, as well as its potential solution strategies, it is also important who actually picks up on the results of R&R on the national level. In other words, in which discursive arena is the gap discussed? Fundamentally, two arenas are relevant: the public and the policy agenda.²² Only if the gap is perceived as substantial in both arenas is a discourse on the topic and subsequent solution strategies (reforms) likely to emerge.

Framing an Issue as Problematic

Participation in international comparative R&R and their influence are in part the product of the convergence of perceptions shared between states regarding the purpose of education policy with respect to citizens and other states. Thus, willing

submission to comparisons with peers may lead to the diffusion of specific practices or modes, which influence the behaviour of actors, especially of those being compared. R&R can establish normative criteria for appropriate behaviour. IOs, such as the OECD, are themselves seen as actors capable of producing shared norms, values and standards.²³ Especially when the membership is comparatively homogenous – as in the OECD, to which only highly developed countries are admitted – do norms and standards receive a high degree of communality.²⁴

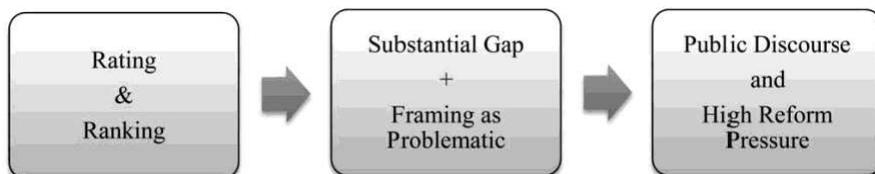
As a precondition for the impact of R&R, it is necessary that the issue under evaluation is nationally perceived and discursively constructed as crucial. Education is generally not framed as an end in itself. It is usually associated with certain superordinate state objectives like micro- and macroeconomic prosperity, technological advancement or even security policy.²⁵ Hence, poor performance in education can anticipate future disadvantages in other relevant domains. R&R in education policy thus have an impact when the issue at stake is framed as challenging another topic of national importance. If R&R successfully reveal and promote the linkage between certain issues, they might re-frame the national understanding of a topic and alter its domestic evaluation. In the case of PISA, the OECD emphasises the connection between educational and economic performance.

The issue of education is nowadays increasingly understood against the background of economic performance, as promoted by the OECD.²⁶ Poor PISA results indicating a national education system's low quality are equated to a risk to the nation's economic prosperity. The need for improving education quality can thus be seen as imperative in order to support national economic performance. By framing education as important for a country's future economic development, the OECD manages to imply the significance of lagging behind other countries. The influence of the OECD particularly stems from its capability to exert normative power and, in this regard, define what is considered as appropriate behaviour.²⁷ Through PISA education was put on the policy agenda of national policy entrepreneurs who further promoted education as the solution for social and economic problems.

Summing up, an issue that receives a performance evaluation contrary to the expected results also has to be defined as problematic in order to trigger national reactions and possible reforms (see Figure 1). That is to say, if in an international comparison the issue X of country A is ranked below its own expectations in the league table, it only causes high pressure for improvements if a negative gap between rank and self-impression exists, and, at the same time, issue X is defined as a crucial factor of country

A. R&R by themselves are not automatically able to exert influence on national

FIGURE 1
THE IMPACT OF R & R



policymaking. Instead, we argue that the mechanisms of self-perception *and* framing substantially shape if and how R&R are received at the national level.

NATIONAL (MEDIA) REACTIONS TO PISA

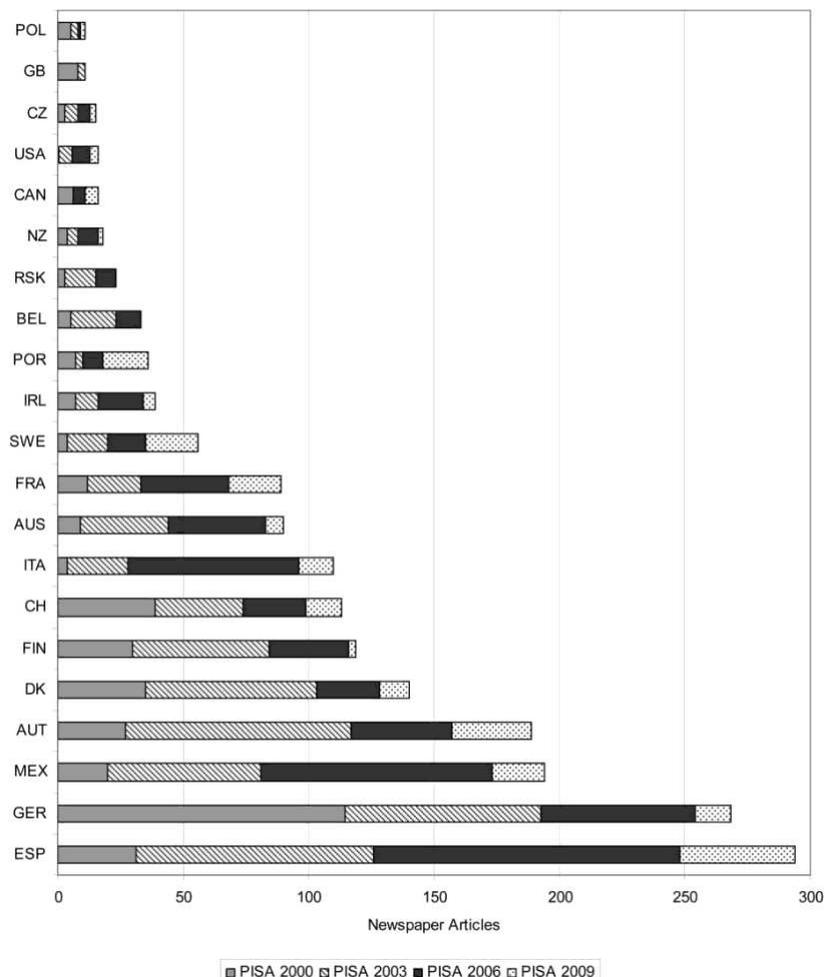
As a rather new phenomenon, only a few studies exist so far evaluating PISA's impact on national education systems, and most of them are case studies.²⁸ In the absence of access to an overall picture of PISA's effects on policies and politics, we use the media attention PISA receives in participating countries as an indicator to analyse the domestic impact the study may have on the public agenda. Obviously, high media response is no guarantee for political reform since the issues are not necessarily on the policy agenda as well. However, the saliency of an issue is commonly reflected by its perception in the media, which displays the topics that are the subject of public controversy.²⁹ If there are many articles in national newspapers dealing with an issue, it is a sign that it is a particularly salient topic in the current national discourse and on the public agenda.

Public reactions to PISA in diverse national contexts reveal differing degrees of saliency of the topic of education: the media coverage in the countries participating in PISA is as diverse as the results. In some countries, such as the US, Canada, Poland and Great Britain, the media has paid comparatively little attention to the OECD study and the respective country's performance in it. In other countries, such as Germany, Spain, Mexico and Austria, by contrast, PISA has been the subject of broad media coverage (see Figure 2), and the publication of the results has been accompanied by an extensive public discourse on education.

Intuitively, a strong positive correlation between the position of a participating country in the ranking and media feedback can be expected. A U-curve of reactions is likely: on the one hand, the empirical evidence of lagging behind other OECD countries in matters of education provides a trigger for public debate; on the other hand, particularly good results can also be expected to receive broad coverage in the national media which reveals how well the national education system and its students are doing. But these arguments do not hold true empirically: a closer look shows that countries that are 'neighbours' in the PISA league table are not inevitably interchangeable with respect to their media reception (see Figure 3). In regard to particularly poor results, the likelihood of it becoming a substantial issue of public discourse increased in countries such as Germany, Spain, Austria and Mexico. However, other countries at the lower end of the PISA league, like the US, Poland and Portugal, did not experience a public outcry reflected in the media coverage. As an overall trend, the worse the rank, the more differentiated the reactions are.

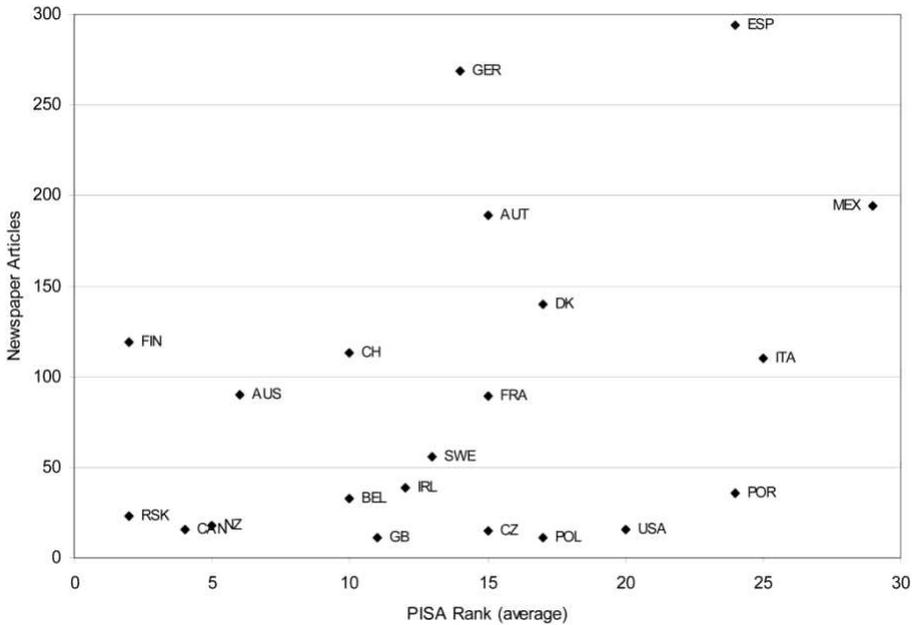
Hence, R&R alone seem to have only a restricted influence on the extent of the public debate on the issue, thus other mechanisms must account for explaining the relationship between a country's PISA results and the reactions in the national media. For the purposes of this article, we explore two extreme cases of media and political reactions to the OECD's PISA study in more depth: Germany and the US. Both countries ranked on the lower end of the PISA league table – whereas Germany was shocked by its PISA results, subsequently launching a public debate about secondary education reforms, reactions were very muted in the US, where neither a public debate nor any related reform attempt was triggered. Since both countries ranked similarly,

FIGURE 2
PISA-RELATED MEDIA COVERAGE



Notes: To assess PISA-related media coverage in each of the 21 OECD countries, one national high-quality daily newspaper with high circulation was analysed for the period from December 2001 to February 2011. In order to be able to draw a dense and more reliable picture of national reaction in Germany and the US we added a second newspaper to our analysis and calculated the mean average. We used primarily newspapers that are available via the online media database 'factiva' (<http://global.factiva.com>). In cases where the data was not continuously accessible through 'factiva', we additionally used the online archives of the respective newspapers. That was the case for the *Frankfurter Allgemeine Zeitung* (2001–11), *Reforma* (Dec. 2001–April 2004), *Irish Independent* (Dec. 2001–Dec. 2002), and *Svenska Dagbladet* (Nov. 2006–Feb. 2011). Articles dealing with the keywords 'OECD' and 'PISA' or 'Programme for International Student Assessment' in their respective national translations were counted. The countries and their national quality newspapers analysed (in parentheses) were: Australia (*The Australian*), Austria (*Die Presse*), Belgium (*De Standaard*), Canada (*National Post*), Czech Republic (*Mladá fronta Dnes*), Denmark (*Politiken*), Finland (*Suomen Tietotoimisto*), France (*Le Monde*), Germany (*Süddeutsche Zeitung* and *Frankfurter Allgemeine Zeitung*), Great Britain (*The Times*), Ireland (*Irish Independent*), Italy (*Corriere della Sera*), Mexico (*Reforma*), New Zealand (*The Press*), Poland (*Rzeczpospolita*), Portugal (*Jornal de Notícias*), Republic of South Korea (*The Korea Herald*), Spain (*El País*), Sweden (*Svenska Dagbladet*), Switzerland (*Neue Zürcher Zeitung*), USA (*New York Times* and *Washington Post*).

FIGURE 3
PISA RANKING RELATED TO MEDIA COVERAGE



Notes: For each country, the average rank of PISA 2000, 2003, 2006 and 2009 is displayed. Additionally, in each of the four circles, the average of the three sub-disciplines (reading, mathematics and science) was assessed.

the sheer results in the measurement as such are not the decisive matter for national reactions; what counts is how the results are perceived.³⁰

TWO SIDES OF THE SAME (PISA) COIN – ASSESSING THE REACTIONS OF GERMANY AND THE US

In this section, we expand our focus beyond the public attention to look at the political reforms in education, comparing the pre-PISA and the post-PISA eras and, hence, include the policy agenda. While Germany significantly reformed its secondary education policy due to its negative PISA performance and the resulting public discourse, the US did not initially show strong reactions to PISA, in either the public or the political sphere.

Germany – The Rediscovery of Education Policymaking Due to the PISA Shock

In Germany, the publication of the first PISA results in late 2001 led to the oft-cited ‘shock’ that pinpointed tremendous deficits in its education performance.³¹ The very essence of the country’s self-perception as a leader of the industrialised world in education was challenged.³² Instantly education became one of the most frequently discussed themes in Germany’s politics and society and became relevant for winning

or losing elections. OECD testing items have since received high coverage in the media; the term ‘PISA’ itself has become a synonym for testing, ranking and rating. Nonetheless, the intensity and extent of the new education debate in Germany after PISA was astonishing. Even though international comparative studies in the mid-1990s, such as Trends in International Mathematics and Science Study (TIMSS), had already sensitised at least the elites in education policymaking to the mediocrity of German students,³³ the broader public remained unaware of the situation. In other words, education became a topic on the policy agenda but remained unnoticed in the public agenda. Unlike previous comparative studies, the publication of PISA ‘led to a public outcry in Germany’.³⁴ The media picked up the PISA report and subjected it to a general debate about the German education system itself and the need for far-reaching reforms. With PISA, the topic of education made it onto the public agenda and at the same time onto the policy agenda, reinforcing the need for policymakers to counteract the performance deficits. While PISA remained a continuously discussed topic, it gained particularly extensive public attention each time new PISA data was published.

Ultimately, by contrasting the international perception of Germany with its self-perception, PISA disclosed the gap to the leading education nations. PISA proved that the outcome of the German education system did not meet the criteria of effectiveness and efficiency. In this regard, it constituted a substantial gap. The first PISA report of 2000 showed that the performances of German students were significantly below the OECD average in all tested areas.³⁵ Overall, Germany was ranked only 20th out of 27 participating OECD countries in PISA 2000. In addition to the poor results, PISA revealed that Germany is one of the OECD member states with the highest level of performance variation across different student groups: in no other industrialised country is academic success so strongly determined by a child’s socio-economic or migration background as in Germany.³⁶ Compared to others, the German secondary education system seems to be less able to reduce existing social inequality.³⁷ Although Germany steadily improved in the consecutive PISA studies – it was ranked 15th in 2003, 11th in 2006 and even ninth in 2009 – educational performance is still regarded as a highly salient issue since the improvement is not seen as sufficient to meet its own expectations. Furthermore, the PISA results of the individual Länder became an increasingly important topic in the German education discourse. The PISA R&R provided data for comparing educational performance at the level of the Länder and showed that some are doing much better than others.³⁸ In this regard, PISA also introduced benchmarks for assessing the educational leaders and laggards within Germany.

Germany was so deeply affected by PISA because this R&R disclosed the standstill of education reforms, which had long characterised German policymaking in this field. While the country underwent far-reaching reforms in the late 1960s and early 1970s, secondary education policy became marginalised in the decades afterwards. In 1964, when a national study by Georg Picht (‘The German Education Catastrophe’) revealed huge deficits in the German education system, education became highly politicised and extensive reform projects followed.³⁹ Most importantly, these reforms aimed at preparing Germany for the challenges of the emerging ‘knowledge society’ by including the demands of the middle class and creating equal opportunities for students irrespective of their social background.⁴⁰ However, since the mid-1970s, no major reform project

had been introduced by Germany. The predominance of other policy issues and political struggles between the left and the right marginalised education policy to the extent that it became an issue almost exclusively dealt with by experts.

Moreover, before PISA, Germany's self-perception of being at the top in educational performance had never been challenged by comparative measures. From the late 1970s to the mid-1990s, Germany did not participate in any international evaluations that could have indicated any deficits in education policy compared to its peers. As a result, Germany perpetuated its educational self-image through a historical understanding, which regarded its educational capability as outstanding. Hence, when PISA 2000 demonstrated Germany's laggard status in education performance, policy-makers were inevitably faced with the shortcomings of the education system compared to other OECD countries. Public demand for the rapid improvement of the quality of secondary education rose. As a consequence, nearly all aspects of secondary education were closely evaluated: school structures, learning processes, and also the education leitmotifs. Even the 'sacred cow' of the German educational system, its three-tiered school system, was scrutinised, questioned and, in some of the Länder – *qua* Basic Law responsible for matters of education in the federation – fundamentally reformed. Since some of the German Länder performed better than others in PISA, the laggard Länder were consequentially under more reform pressure. The new century heralded a new enthusiasm for educational reform through the introduction of PISA 2000.

PISA ultimately provided for Germany's willingness to reform. As early as December 2001, within weeks after the release of PISA 2000, the reform efforts by the responsible German education policymakers at the federative level, the Standing Conference of the Ministers of Education and Cultural Affairs (KMK) as the main forum to coordinate education policy in Germany, included measures to improve early education, assist underprivileged students, provide for better teacher training and extend day care.⁴¹ In order to increase the quality of education, in 2002 the KMK agreed upon the introduction of educational standards and the establishment of a central agency for monitoring compliance with these standards.⁴² The whole catalogue of reform measures can be directly linked to PISA and the identified deficits of the German education system. However, these proposed reform measures were not 'invented' in the few days after the publication of PISA; in fact, they were already-existing compiled reform schemes. Yet, it was only through PISA that they made it onto the political agenda and were accepted.⁴³

Through Germany's poor PISA results, education was increasingly framed as crucial for future national economic prosperity. In this regard, the liberal education tradition in Germany was challenged by the ostensible economic attitude PISA promoted.⁴⁴ Whereas prior to PISA education was first and foremost seen as a microeconomic issue emphasising the training of the individual for a profession and as a citizenship issue preparing them for participation in the democratic society; the post-PISA debate is increasingly characterised by a shifted emphasis toward education's macroeconomic dimensions.⁴⁵ Education is now regarded as central to the economic performance of a whole country: poor performance in education might also carry the danger of jeopardising Germany's wealth. Understood in this way, the poor PISA result of a considerable ratio of German students in turn puts the future German economic condition at risk.

In sum, through the double-edged process of evaluating the national system and comparing it to other countries, PISA shook the German self-perception to its core and created an immense pressure on German policymakers to improve the secondary education system. As Kingdon put it, a problem window was opened that presses in on the government.⁴⁶ PISA not only successfully promoted the OECD's economic centred ideas and recommendations about education but also influenced Germany at the level of policymaking to adapt instruments for assessing school performance. Within a very short period of time, the country established a culture of comparison in education policymaking. The reforms implemented in Germany's education system reflect a paradigm shift that comprises the enhanced orientation towards an output perspective on education, as promoted by the PISA study. The already existing but rudimentary debates about quality development, quality assurance, teachers' education and curricular issues were increasingly fostered by PISA. Today, education policymaking does not simply focus on the fulfilment of certain principles but instead integrates the new governance mode of 'evidence-based policymaking'.

The US – How a Country (Almost) Disregards its PISA Results

Unlike in Germany, the first three cycles of PISA and their results have remained virtually unheard of in the US. The general public and the media did not take particular notice of this international R&R or its results, despite the fact that the US performed consistently below or on average among participating countries. PISA did not lead to a shock as in Germany – despite its similarly poor results. It was only after the fourth PISA study was published in 2009 that the American public became at least modestly aware of this international R&R. Unlike the previous PISA studies, the American results were framed as problematic for its national economy after the latest study.

With the first three PISA rounds, the US ranked continuously worse. It scored 17th in 2000, 21st in 2003 and in 2006 fell to 23rd among the participating OECD member states. In the 2009 PISA study the US slightly improved its ranking, once again becoming 17th. Despite these poor results, PISA is far from being a theme in the American news. In fact, the first publication of the PISA results in 2001 was not covered at all in the two major daily newspapers, the *New York Times* and the *Washington Post*. Compared to Germany, it is safe to say that PISA was not an issue on the US public agenda. The average American does not have any knowledge of PISA as an international comparative education study.

Unlike the German case, PISA did not reveal a substantial gap in the US self-perception of the secondary education system. Thus, the OECD study did not surprise the Americans in regard to their school performances as PISA did for other countries. On the contrary, the American public and policymakers have been aware of its poor secondary education system and the need for reforms for decades. The country has a long history of participating in education assessments, mainly across US states. The testing and accountability movement strongly influenced the performance-based measurement of education achievement since the 1980s.⁴⁷ Therefore, the US had introduced the paradigm of output-oriented education evaluation much earlier than, for instance, Germany and also before PISA started. While PISA did not show any new or surprising

insights on the American education system the study was just one among others and did not make it onto the public and political agenda. From this viewpoint, the US fared equally poorly in PISA as it did in other international comparative studies, such as TIMSS. PISA confirms that the American secondary school system needs continuous improvement.

In fact, education policy has been an issue on the public as well as the political agenda from as early as 1957, when the Soviet Union launched the first satellite. From the American viewpoint, Sputnik demonstrated the Soviet Union's technological superiority over the US or, at the very least, its equality in the field. The reasons for this American 'lag' were primarily seen as rooted in the education system, which excluded too many people from participating in social progress. The Sputnik shock led to the first comprehensive reform of the American education system, and diverse programmes were set up with the intention of improving the curriculum, particularly in the natural sciences.⁴⁸

As a synonym for traumatic education failure, Sputnik has plagued the US perception of education policy ever since. When the US government enlisted a national commission to conduct a study of the American educational system 25 years after Sputnik, it intended to review the impressive progress made. However, the resulting 1983 report, 'A Nation at Risk: Imperatives for Educational Reforms', presented devastating outcomes: with 23 million adults and 17 per cent of juveniles illiterate, the American educational system was in abysmal shape. The report – issued at the height of Cold War during the Reagan years – once again triggered broad public concern and a series of political educational reforms aimed at creating standards, providing for accountability, enabling school choice and improving the quality of teaching.⁴⁹

Similar to the conceptualisation of PISA, consecutive reforms following 'A Nation at Risk' have propagated outcome-oriented types of educational teaching. Continuous reform efforts across and within US states for improving the secondary school system have been on the political agenda ever since.⁵⁰ In particular, the education act of 2001, 'No Child Left Behind Act (NCLB)', reauthorised a number of federal programmes that aimed to improve the performance of US primary and secondary schools by increasing the standards of accountability for states, school districts and schools as well as providing parents more flexibility in choosing the schools their children will attend.⁵¹ The NCLB requires states wishing to receive federal funding for schools to develop assessments in basic skills that are to be conveyed to all students in certain grades. However, the act did not establish a national performance standard. On the contrary, each federal state can – in accordance with the principle of school autonomy – set its own standards.⁵² Thus, within the domestic sphere, the US has increasingly aimed at standardisation and comparative evaluations over the last two decades, but it did so without reference to international R&R, such as PISA.

However, political elites, think tank representatives, lobby groups and interest groups are well aware of PISA and the particular deficits this study revealed about the US secondary school system, despite the fact that international comparisons, such as PISA, have remained almost unnoticed among the broader public. Those involved in education policymaking in the US mainly use the information PISA delivers to extract *best practice* examples from abroad revealed by the study. But unlike the German case, PISA is just one 'product' alongside other national and international

comparative studies used for political purposes; it is not the reference or catalyst for change.

Part of the reason why international R&R do not receive nationwide responses in the US can be traced to the country's highly decentralised and heterogeneous education system. Institutionally, the American government has only a small voice in the education realm, as the US Constitution does not provide for any rights to the federal government in the field of education; this is delegated to the state level. Before 1979, the American Federal Ministry of Education did not even exist. As a result, from an American point of view *international* comparisons (and the conclusions drawn from them) are often seen as inapplicable to the local US context. Therefore, R&R also cannot touch per se any 'American' self-perception of education as a whole.

However, unlike in Germany, current reforms are undertaken without direct reference to PISA. Although the OECD has tried to make PISA better known in the US,⁵³ the 'Race to the Top' (R2T) programme as part of the American Recovery and Reinvestment Act of 2009 or the reauthorised education act of March 2010, 'A Blueprint for Reforms', do not mention the US PISA results, even though major goals are parallel to what PISA propagates.⁵⁴ The R2T, for example, promotes better outcome-oriented measures of student performance. The blueprint favours a school ratings system based on the pass-fail scheme being replaced with a system measuring individual student progress and school evaluation. The blueprint also demands clear intervention in failing schools but reduces government intervention in well-administered middle-ranking schools. In particular, the new measurements under the Obama government include funds for preschool programmes, merit pay for teachers and the establishment of more charter schools. Thus, although many goals of PISA and the Obama government are similar, PISA is framed on domestic grounds rather than being oriented on the OECD's paradigm.

Only the latest publication of PISA results in December 2010 triggered some moderate reactions in the broader public and the political realm since the American PISA results were framed as problematic for the national economy. What was decisive for the US to become aware of PISA was not so much the continuously mediocre results of its own students, but rather the outstanding test scores of Shanghai school students tested for the first time. This made the US media more aware of high school standards in other countries. The sensational results of the most direct opponent for economic markets provoked reactions on the US side. Education experts as well as President Obama himself made the analogy with Sputnik when interpreting the PISA 2009 results of the US.⁵⁵ The Chinese result triggered a new wave of all-encompassing education reforms in the US, in an attempt to make the country more competitive with regard to its economic performance. Thus, PISA now provided a policy window to frame secondary education as problematic, placing it on the public as well as on the political agenda.

In brief, PISA received only minimal reaction in the US compared to other participating states. The country's poor PISA scores did not present a gap between the American self-image concerning the performance of its secondary education system and the actual results in the OECD ranking. Thus, PISA did not open a problem window to which government had to respond. The shortcomings of the American school system have ultimately become an undisputed matter of fact, with the consistent

implementation of numerous reform processes and new strategies. The PISA results are known to an elite group of policymakers and education groups in the US, but they constitute only one of a variety of national and international tests in which the country takes part. Despite the reforms, schools in the USA have only improved to a limited extent. Whereas domestic R&R within the US play a decisive role today concerning school choice, the international standing of the US as a whole generally does not receive much attention. Only recently, however, have the comparatively poor secondary school education results of the US been framed as problematic, since it is recognised that its economy, and thus the country's prosperity is being challenged.

CONCLUSION AND OUTLOOK

In order to explain differences in reactions of national media coverage and follow-up political reforms to PISA, it is insufficient to simply refer to a nation's actual position in the ranking. As our media analysis has shown, there is no obvious correlation between the rank of a state and its reaction: states that are evaluated similarly vary substantially regarding their national responses. In other words, a poor ranking is not sufficient to trigger a national education debate and initiate reforms. Therefore, to explain the power international R&R can exert, we need an approach that goes beyond a simple functional rank-reaction model. In this article, we conceptualised R&R with respect to a theoretical approach of governance in combination with agenda-setting theory. We argued that two dimensions need to be considered when seeking to understand under what conditions R&R have an impact on national policymaking, namely that the relation between the self-perception and the actual results needs to be taken into account, and the issue – in our case, education policy – needs to be framed as problematic within the national context.

In the empirical part, we applied our theoretical concept to two countries which were on average ranked similarly poorly in the OECD's PISA study but which reacted differently: Germany with a broad public debate about education and comprehensive reform efforts and the US with very modest media response and almost no political reaction. Our analysis showed that Germany witnessed a substantial gap between its PISA results and its self-image as an education nation (*perception*) in the time of a knowledge-based society in which education became a significant 'resource' for the economic well-being of the country (*framing*). In particular, the poor results uncovered problematic aspects of the German education system such as fostering social inequality and system inefficiency in integrating children of migrant families. In contrast, in the case of the US, PISA has only lately begun to trigger some debate about the American education system due to the comparatively high results of China, its direct economic rival (*framing*), but it did not surprise the country in regard to the general low standards of the education system (*perception*). Our theoretical framework presents an attempt to provide an explanation for the puzzling different responses to comparative rankings. Beyond the actual effects and non-effects PISA exerted on both states, our study provides insights into the authority an expert organisation, such as the OECD, can have with R&R. Even if the OECD is assigned the task of producing information to guide state-based decision-making, the process of carrying out the actual implementation remains in the hands of the experts working for the IO. From a more general

point of view, the case study of the OECD and its PISA study shows that IOs that generate and publish comparative evaluations gain power by setting standards.⁵⁶ Furthermore, IOs can contribute to the definition of a certain issue as a problem to be dealt with. PISA is an example of the variation in the reception of international stimuli at the national level. As shown in the case of the US, PISA more or less stated the obvious, and thus its technique of governance by comparison was not perceived as a groundbreaking revelation of new phenomena; in contrast, in Germany the comparison told a ‘new story’ and provided essential insights. As a precondition, an IO needs to be perceived as a legitimate authority and expert organisation in order to be successful in influencing policymaking at the nation-state level.

Further research needs to evaluate our findings against additional arguments. Both states possess different political cultures as regards their reception of IOs and their governance. The US perceives itself as independent and does not feel bound to international commitments as much as other states do. From this point of view, it is not surprising that the American secondary education sector is only slightly affected by the PISA survey conducted by the OECD. In addition, the more competences for education are located at the local or private level, the less international initiatives should bother the country as a whole. While in Germany state authorities (at the level of the *Länder*) are responsible for education policymaking, the US system is organised much more locally and is more privatised. PISA, however, assesses national education performance, and as a result it is difficult to clearly translate poor results to the local level in the US. In contrast, poor PISA results of German students impacted on the education system as a whole. It remains to be seen to what extent international R&R can become more influential in the US as education policy becomes more centralised.⁵⁷

Our analysis of the German and the US reactions to the OECD’s PISA study can, however, only be a first step in grasping the influence that governance by comparison may have. It would be a promising endeavour to compare respective reactions to PISA in other OECD countries in terms of the concepts of self-perception and framing. Expanding the scope of governance by comparison to other policy fields beyond education would be useful in clarifying to what extent the power of the OECD (or other IOs) depends on its authority and status in the international community.

ABOUT THE AUTHORS

Kerstin Martens is Associate Professor of International Relations at the University of Bremen. She heads the 12-year research project on ‘Internationalisation of Education Policy’ at the Collaborative Research Centre ‘Transformations of the State’ at the University of Bremen.

Dennis Niemann is Research Assistant in the project on ‘Internationalisation of Education Policy’ at the Collaborative Research Centre ‘Transformations of the State’ at the University of Bremen.

NOTES

1. Prominent examples of current influential R&R of states are the international credit rating agencies and their prognoses of the liquidity and credit worthiness of states during the global financial crisis.

2. Available from <http://www.arwu.org/ARWU2010.jsp> (accessed 24 Aug. 2012).
3. S. Dillon, 'Top Test Scores From Shanghai Stun Educators', *New York Times*, 7 Dec. 2010, http://www.nytimes.com/2010/12/07/education/07education.html?_r=1 (accessed 24 Aug. 2012).
4. Available from http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1,00.html (accessed 23 Aug. 2012).
5. M. Henry, B. Lingard, F. Rizvi and S. Taylor, *The OECD, Globalisation and Education Policy* (Oxford: Pergamon, 2001); S. Grek, 'Governing by Numbers: the PISA "Effect" in Europe', *Journal of Education Policy* 24/1 (2009), p.28; Organisation for Co-operation and Development (OECD), *Work on Education* (Paris: OECD Publications, 2010).
6. For further information on PISA, see <http://www.pisa.oecd.org>.
7. R. Rinne, 'Too Eager to Comply? OECD Education Policies and the Finnish Response', *European Educational Research Journal* 3/2 (2004), pp.454–85.
8. B. Reinalda and B. Verbeek (eds), *Autonomous Policy Making by International Organizations* (London: Routledge, 1998); M. Barnett and M. Finnemore, *Rules for the World. International Organizations in Global Politics* (Ithaca, NY: Cornell University Press, 2004); D.G. Hawkins, D.A. Lake, D. Nielson and M.J. Tierney (eds), *Delegation and Agency in International Organizations* (New York: Cambridge University Press, 2006); J. Joachim, B. Reinalda and B. Verbeek (eds), *International Organizations and Implementations* (New York: Routledge, 2008); K. Martens and A.P. Jakobi (eds), *Mechanisms of OECD Governance – International Incentives for National Policy Making?* (Oxford: Oxford University Press, 2010).
9. T.M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, NJ: Princeton University Press, 1995); P. Miller, 'Governing by Numbers: Why Calculative Practices Matter', *Social Research* 68/2 (2001), pp.379–96; W.N. Espeland and M. Sauder, 'Rankings and Reactivity: How Public Measures Recreate Social Worlds', *American Journal of Sociology* 113/1 (2007), pp.1–40; F.A. van Vught and F. Ziegele (eds.), *Multidimensional Ranking. The Design and development of U-Multirank* (Dordrecht: Springer, 2012).
10. Grek, 'Governing by Numbers'; K. Martens, 'How to Become an Influential Actor – the "Comparative Turn" in OECD Education Policy', in K. Martens, A. Rusconi, and K. Leuze (eds), *New Arenas of Education Governance – The Impact of International Organizations and Markets on Education Policy Making* (Houndmills: Palgrave Macmillan, 2007), pp.40–56; T. Jakobi and W. Lamping, 'Die Macht der Rangliste? Die Reaktionen auf internationale Leistungsvergleiche im deutschen Politikprozess', *Zeitschrift für Politikwissenschaft* 21/3 (2011), pp.365–89.
11. J. Gerring, *Case Study Research: Principles and Practices* (New York: Cambridge University Press, 2006).
12. Martens and Jakobi, *Mechanisms of OECD Governance*.
13. The German and the US secondary education systems also differ in various respects, e.g. the US system has a far stronger private sector than the German system and also recognises forms of school selection other than the one in the Germany system, such as school vouchers. Nevertheless, systemically, in both systems the sub-national level (the states in the US and the Länder in Germany) is in charge of education. However, both systems have also institutionalised coordinating bodies for national education policies. M. Busemeyer, *Die Bildungsausgaben der USA im internationalen Vergleich: Politische Geschichte, Debatten und Erklärungsansätze* (Wiesbaden: Deutscher Universitäts-Verlag, 2006), ch.2; P.J. McGuinn, *No Child Left Behind and the Transformation of Federal Education Policy, 1965 – 2005* (Lawrence, KS: University Press of Kansas, 2006).
14. K.W. Abbott, and D. Snidal, 'Hard and Soft Law in International Governance', *International Organization* 54/3 (2000), pp.421–456.
15. K.A. Mingst, 'Global Governance: The American Perspective', in R. Väyrynen (eds), *Globalization and Global Governance*, (Lanham/Oxford: Rowman & Littlefield, 1999), pp.87–104; M. Djelic and K. Sahlin-Andersson, 'Introduction: A World of Governance: The Rise of Transnational Regulation', in M. Djelic and K. Sahlin-Andersson (eds), *Transnational Governance. Institutional Dynamics of Regulation* (New York: Cambridge University Press, 2006), pp.1–28; for education Grek, 'Governing by Numbers'.
16. J.W. Kingdon, *Agendas, Alternatives, and Public Policies*, 2nd edition (New York: HarperCollins, 1995).
17. *Ibid.*, p.3.
18. D. Béland, 'Ideas and Social Policy: An Institutionalist Perspective', *Social Policy & Administration* 39/1 (2005), pp.1–18; Kingdon, *Agendas, Alternatives, and Public Policies*.
19. P. Sahlberg, 'Education Reform for Raising Economic Competitiveness', *Journal of Educational Change* 7/4 (2006), pp.259–87.
20. Abbott and Snidal, 'Hard and Soft Law in International Governance', p.5.
21. Kingdon, *Agendas, Alternatives, and Public Policies*.

22. S. Soroka, *Agenda-setting Dynamics in Canada* (Vancouver: University of British Columbia Press, 2002).
23. M. Finnemore, 'International Organizations as Teachers of Norms: the United Nations Educational, Scientific, and a Cultural Organization and Science Policy', *International Organization* 47/4 (1993), pp.565–97.
24. V. Rittberger and B. Zangl, *International Organization. Polity, Politics and Policies* (New York: Palgrave Macmillan, 2006).
25. K. Martens and A. Weymann, 'The Internationalization of Education Policy – Towards Convergence of National Paths?', in A. Hurrelmann, S. Leibfried, K. Martens and P. Mayer (eds), *Transforming the Golden Age Nation State* (Basingstoke: Palgrave Macmillan, 2007), pp.152 – 72.
26. Henry et al., *The OECD, Globalisation and Education Policy*; Sahlberg, 'Education Reform for Raising Economic Competitiveness'.
27. I. Manners, 'Normative Power Europe: A Contradiction in Terms?' *Journal of Common Market Studies* 40/2 (2002), pp.235 – 58.
28. H. Ertl, 'Educational Standards and the Changing Discourses on Education: The Reception and Consequences of the PISA Study in Germany', *Oxford Review of Education* 32/5 (2006), pp.619–34; K. Tillmann, K. Dederich, D. Kneuper, C. Kuhlmann and I. Nessel, *PISA als bildungspolitisches Ereignis* (Wiesbaden: VS Verlag für Sozialwissenschaften, 2008); Grek, 'Governing by Numbers'; K. Martens, A.-K. Nagel, M. Windzio and A. Weymann (eds), *Transformation of Education Policy* (Houndmills: Palgrave Macmillan, 2010). See also <http://www.knowandpol.eu> for ongoing work on this issue.
29. T.A. Collins and C.A. Cooper, 'Case Salience and Media Coverage of Supreme Court Decisions: Toward a New Measure', *Political Research Quarterly* 65/2 (2012), p.397.
30. W. Van Dooren, 'Performance Indicators: a Wolf in Sheep's Clothing?', Paper presented at the Wissenschaftszentrum Berlin at the Conference on Changing Educational Accountability in Europe; 24 – 25 June 2008, Berlin.
31. Ertl, 'Educational Standards and the Changing Discourses on Education'; C. Lundahl and F. Waldow, 'Standardisation and "Quick Languages": The Shape-Shifting of Standardised Measurement of Pupil Achievement in Sweden and Germany', *Comparative Education* 45/3 (2009), pp.365– 85.
32. F.-O. Radtke, 'Die Erziehungswissenschaft der OECD – Aussichten auf die neue Performanz-Kultur', *Erziehungswissenschaft* 27/14 (2003), pp.109 – 36.
33. R.G. Heinze, 'Die Bildungskatastrophe – Ein alter Hut?', in B. Fahrholz, S. Gabriel and P. Müller (eds), *Nach dem Pisa-Schock. Plädoyer für eine Bildungsreform* (Frankfurt am Main: Hoffmann und Campe, 2002), pp.16–21.
34. A. Ammermüller, 'PISA: What Makes the Difference?', discussion paper DFG-Forschergruppe 04/07 (2004), p.2.
35. J. Baumert, P. Stanat and A. Demmrich, 'PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie', in J. Baumert et al. (eds), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (Opladen: Leske + Budrich, 2001), pp.15–68.
36. H. Kiper and U. Kattmann, 'Basiskompetenzen im Vergleich – Überblick über Ergebnisse der PISA-Studie 2000', in B. Moschner, H. Kiper and U. Kattmann (eds), *PISA 2000 als Herausforderung* (Baltmannsweiler: Schneider Verlag Hohengehren, 2003), pp.15 – 37; Ertl, 'Educational Standards and the Changing Discourses on Education'.
37. H. Loeber and W. Scholz, 'Von der Bildungskatastrophe zum Bildungsschock – Zur Kontinuität sozialer Benachteiligung durch das deutsche Bildungssystem', in B. Moschner, H. Kiper and U. Kattmann (eds), *PISA 2000 als Herausforderung* (Baltmannsweiler: Schneider Verlag Hohengehren, 2003), pp.241–85.
38. J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schmieder, K.-J. Tillmann and M. Weiß. *Pisa 2000 – Ein Differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (Opladen: Leske + Budrich, 2003); J. Allmendinger and S. Leibfried, 'Education and the Welfare State: The Four Worlds of Competence Production', *Journal of European Social Policy* 13 (2003), pp.63 – 81.
39. G. Picht, *Die deutsche Bildungskatastrophe: Analyse und Dokumentation* (Olten and Freiburg: Walter-Verlag, 1964).
40. P. Massing, 'Konjunkturen und Institutionen der Bildungspolitik', in P. Massing (ed.), *Bildungspolitik in der Bundesrepublik Deutschland* (Schwalbach: Wochenschau Verlag, 2003), pp.9 – 55; K. Mayer, 'Das Hochschulwesen', in K. Cortina, J. Baumert, A. Leschinsky, K. Mayer and L. Trommer (eds), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (Reinbek: Rowohlt Taschenbuch Verlag, 2005), pp.581 – 624.
41. MKM and BMBF, *Ergebnisse von PIRLS/IGLU 2006-I und PISA 2006-I: Gemeinsame Empfehlungen der Kultusministerkonferenz und des Bundesministeriums für Bildung und Forschung* (2008), available

- from http://www.bmbf.de/pubRD/gemeinsame_empfehlungen_iglu_und_pisa_06.pdf (accessed 27 Aug. 2012); D. Carey, 'Improving Education Outcomes in Germany', *Economic Department Working Papers, OECD* (2008).
42. W. Nieke, 'Schulreform: Wie reagiert die Fachpolitik auf bildungswissenschaftliche Bestandsaufnahmen?', in T. Hansel (ed.), *PISA – und die Folgen? Die Wirkung von Leistungsvergleichsstudien in der Schule* (Herbolzheim: Centaurus Verlag, 2003), pp.197 – 214.
 43. Tillmann et al., *PISA als bildungspolitisches Ereignis*.
 44. D. Niemann, 'Turn of the Tide – New Horizons in German Education Policy through IO Influence', in K. Martens, A.-K. Nagel, M. Windzio and A. Weymann (eds), *Transformation of Education Policy – The Impact of the Bologna Process and the PISA Study in Comparative Perspective* (Houndmills: Palgrave Macmillan, 2010), pp.77 – 104.
 45. Radtke, 'Die Erziehungswissenschaft der OECD'.
 46. Kingdon, *Agendas, Alternatives, and Public Policies*, p.174.
 47. D. Ravitch, 'A Brief History of Testing and Accountability', *Hoover Digest* 4 (2002), available from <http://www.hoover.org/publications/hoover-digest/article/7286> (accessed 10 Nov. 2012).
 48. D. Ravitch, *The Troubled Crusade: American Education 1945–1980* (New York: Basic Books, 1983).
 49. R. Gross and B. Gross (eds), *The Great School Debate* (New York: Touchstone, 1985).
 50. See for example K.R. Kosar, *Failing Grades. The Federal Politics of Education Standards* (Boulder, CO and London: Lynne Rienner Publisher, 2005); McGuinn, *No Child Left Behind*.
 51. P. McGuinn, 'Swing Issues and Policy Regimes: Federal Education Policy and the Politics of Policy Change', *Journal of Policy History* 18/2 (2006), pp.205 – 40.
 52. Kosar, *Failing Grades*.
 53. Organisation for Co-operation and Development (OECD), *Lessons from PISA for the United States. Strong Performers and Successful Reformers in Education* (Paris: OECD Publishing, 2011).
 54. Available from <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf> (accessed 22 Aug. 2012); <http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf> (accessed 22 Aug. 2012).
 55. Dillon, 'Top Test Scores From Shanghai Stun Educators'.
 56. Martens and Jakobi, *Mechanisms of OECD Governance*.
 57. E.H. Debray, *Politics, Ideology and Education. Federal Policy During the Clinton and Bush Administrations* (New York: Columbia University Press, 2006).