

Dixon et al. 2013, published online in *Public Administration*
<http://onlinelibrary.wiley.com/doi/10.1111/padm.12013/abstract>

A Lever for Improvement or a Magnet for Blame? Press and Political Responses to International Educational Rankings in Four EU Countries.

Ruth Dixon¹, Christiane Arndt², Manuel Mullers², Jarmo Vakkuri³, Kristiina Engblom-Pelkkala³, and Christopher Hood¹

1. Department of Politics and International Relations, University of Oxford, UK.
2. Affiliated to the Graduate School of Governance, Maastricht University, The Netherlands.
3. School of Management, University of Tampere, Finland

Corresponding author:
Dr Ruth Dixon (ruth.dixon@politics.ox.ac.uk)

Acknowledgements

We acknowledge the support of the ESRC Public Services Programme, All Souls College, Oxford, and the Graduate School of Governance at Maastricht University. We are grateful to Moritz Neuffer for help with the German literature analysis, and to Hubert Ertl (Oxford University), Deborah Wilson (Bristol University), Pablo Zoido (OECD Education Directorate) and three anonymous reviewers for their comments on previous drafts of this article.

Abstract

Educational performance rankings elicit extensive press coverage and varied political responses. To investigate how the negativity of press coverage was related to the rankings results and political response, we compared the domestic press coverage of two educational rankings (the 2006 editions of the OECD's Progress in Student Achievement (PISA) and the IEA's Progress in International Reading Literacy Study (PIRLS)) in four European countries (Germany, Finland, France, and Britain). We found negative press coverage in all four countries. Strikingly, Finland had the same percentage of negative articles as Germany, despite much better performance. British politicians received far more personal blame for poor results than their French or German counterparts, but the political response was strongest in Germany and Britain. We discuss our findings in the context of the effects such rankings might be expected to have, and argue that these rankings alone do not provide sufficient levers to improve educational performance.

Introduction:

International rankings are increasingly used by national and international organizations and policy-makers to evaluate the success or failure of policies, to justify or propose reforms and to allocate funding. The growth of international rankings of governance and public services have been much commented on, particularly from the perspective of the validity and reliability of the information provided by such rankings (e.g. Besançon 2003; Rotberg and West 2004; van de Walle 2006; Hood et al. 2008; Arndt 2008). And there is an extensive literature on the reactions

of institutions to being ranked (e.g. Espeland and Sauder 2007; Chatterji and Toffel 2010; Frederickson and Stazyk 2010). Little though, has been written on the domestic reception of such rankings by the media, and particularly on cross-national comparisons of such coverage, despite the fact that the publication of international league tables often excites considerable press coverage, both in countries that do well and those that do poorly (Schmidt 2003a, 2003b, 2003c; Gruber 2006).

In this study, we compared the domestic press and political reactions in four EU countries to two international educational surveys that were held in 2006 and published in late 2007, namely Progress in Student Achievement (PISA) and Progress in International Reading Literacy Study (PIRLS). This ambitious and unique cross-national survey allowed us to analyse the political response to rankings in the context of the tenor of the press coverage and the level of blame in the articles. We discuss these results in the light of the literature on rankings, negativity and blame-avoidance.

There are previous reports of negativity in press treatment of domestic public service performance targets in England (James 2004; Hood and Dixon 2010) and observations of asymmetric electoral payoffs for politicians for meeting or failing to meet such targets at local government level in England in the 2000s (see James and John 2007; Boyne et al. 2010). These earlier findings for domestic health, education and local government performance measurement systems in England in the early to mid 2000s might possibly be thought to reflect factors specific to England, such as its exceptionally centralised political system (such that the delivery of public service performance is seen as something that national politicians are responsible for (Hood 2007; Barber 2007)), its greater reliance on measured performance indicators compared to other EU countries (see Harrison and Pollitt 1992), and the markedly challenging tone of its national (especially tabloid) newspapers. They might also be thought to reflect factors specific to the politics of England in the period of Tony Blair's second electoral term, such as disillusion with the New Labour spin machine and (after 2003) increased distrust of government among the commentariat associated with the Iraq war and occupation in 2003.

This paper aims to explore how generalizable the 'negativity' phenomenon is by exploring domestic press reactions to national performance data reported in international educational rankings in Finland, France, Germany and Britain. By negativity, we mean the preponderance of negative comment towards the home country's results or its education system, as judged by content analysis of a sample of press articles. This leads us to the first two questions addressed in this study: was negativity indeed universal for the cases and countries examined? (See hypothesis 1, below.) And, if so, was there evidence of negativity *bias*, that is, did the negativity actually reflect poor performance, or was it found regardless of performance (hypothesis 2)?

Negativity bias can be found in many contexts (Baumeister et al. 2001; Rozin and Royzman 2001; Lau 1982; Soroka 2006) and in the asymmetric payoffs that politicians experience, suffering far worse electoral punishment for failure than they receive electoral credit for success (Weaver 1986; James and John 2007) as mentioned above. As a result, politicians have been claimed to have a general tendency to adopt blame-avoidance strategies (Weaver 1986; James 2004; Hood 2010) and respond to and seek to manipulate levels of blame (Hood et al. 2009).

Our third question is therefore: by participating in international rankings, do politicians expose themselves to the risk of being personally blamed (hypothesis 3)?

A further question concerns how international rankings motivate institutions to change their behaviour as against other kinds of rankings. Are PISA and PIRLS similar to rankings of (for instance) US law schools (Sauder 2008; Sauder and Espeland 2009), other graduate programmes (Frederickson and Stazyk 2010) and private businesses (Chatterji and Toffel 2010)? This literature documents the 'reactivity' of institutions (Espeland and Sauder 2007), that is, radical changes in behaviour that occur in response to publicly available ranking information. Deans and faculty are hired or fired; publicity is sent out by the ranked institutions to the rankers; and funding of departments depends on the position in the rankings. Proponents of such rankings suggest that the reputational advantage in attracting students or customers drives up standards, though critics point to the distorting effects on the overall performance of the institution. Do international educational rankings have such far-reaching effects, and if so, on whom?

To explore this question, we should consider the aim of such rankings. If the ranking is a lever for improvement of educational standards, how is this lever meant to work? Unlike the rankings of universities or private companies described above, there is no direct link at the level of the delivery organization (the school), no-one's job is on the line, and there is no direct incentive for any individual school to change as the results of participating schools are not published. So the effect of PISA or PIRLS presumably rests on making 'the government' feel responsible for national performance via the credit and/or blame arising from the results. This would lead government ministers to put pressure on the school system, using the results as justification. The fact that ministers receive the results mediated by the press is also likely to be a significant factor in determining their response. We might expect that the way in which the results are reported will influence the political response for the following reasons.

Oliver James (2011) has shown that information about public service performance affects citizens' expectations in an asymmetric way (negative information has a relatively greater negative effect than positive information has in the opposite direction). James and John (2007) have shown that negative performance information affects voting intention, in that politicians are disproportionately penalized for poor performance. Stuart Soroka (2006) found that negative media coverage of economic performance enhanced the effect of negative economic performance itself on public opinion of the economy. Soroka argued that the effect of media coverage may be even more pronounced where the public has little direct experience of the performance in question. Since the PISA and PIRLS results are only readily available to the public via the media (unless the citizen is motivated to study the organizations' websites), few people will have other ways to learn about the relative performance of their country's education system. It is therefore likely that negative press coverage of these rankings will have a strong negative effect on public opinion of a country's education policies. So to the extent that politicians care about public opinion and hence voting intention, they are likely to respond to negative press coverage of the results correspondingly more urgently than they would respond if the results were only released via a website or to education professionals. This gives rise to our hypothesis 4, that the intensity and tenor of the press coverage will influence the political response.

The hypotheses this study aims to test are therefore:

(1) *Negativity is found in press reports of educational rankings.*

And, if hypothesis (1) was found to be true,

(2) *Negativity bias is found in press reports of educational rankings.*

(3) *Politicians are personally blamed in press reports of educational rankings*

(4) *Educational rankings will elicit a political response that is influenced by the intensity and tenor of the press coverage*

We selected Germany, Finland, France and Britain as our cases as they span a range of opinions and attitudes to education and the measurement of educational performance, as well as having a wide range of performance in the educational rankings selected for our study (PISA and PIRLS). To summarise the PISA 2006 results, Finland's performance was 'excellent' being ranked first or second in the three subject areas (maths, reading and science), Germany was 'improving' (compared to its 2000 and 2003 results), and Britain and France 'deteriorating'. In PIRLS, Germany's results were stable, and France and Britain's were deteriorating (Finland did not take part in PIRLS). The performance results are shown in Table 1. These four EU democracies have different state traditions, political cultures and governmental structures. They each have a well established press, though as discussed later, their different journalistic traditions produce certain differences in their press cultures. We chose six newspapers from each country that reflected as far as possible a range of political, regional and cultural outlooks, as well as each representing a sizable circulation.

The PISA and PIRLS educational rankings

We selected the PISA and PIRLS studies as being international rankings that attract a high level of press coverage and are among the most respected of international comparisons of public services. PISA assesses 15-year-old school students in mathematics, science and reading literacy on a three-yearly basis and has been run by the Organisation for Economic Cooperation and Development (OECD) since 2000. PIRLS (assessing reading literacy in 10-year-olds) has been run by the International Association for the Evaluation of Educational Achievement (IEA) since 2001. These programmes developed from international testing programmes that began in the 1950s, and have similar methodologies (Hutchison and Schagen 2007). In 1993, the OECD produced its first survey of international educational statistics known as *Education at a Glance*. Since 2006 the PISA results have been the only educational outcome indicators reported in that publication. PISA 2006 involved 57 countries and almost 400,000 pupils worldwide. PIRLS 2006 covered 41 countries and about 200,000 pupils.

Table 1 about here

Table 1. Performance in PISA and PIRLS.

Previous studies of press coverage of PISA

The extensive press coverage of the previous (2000 and 2003) editions of PISA has been described elsewhere (e.g. Schmidt 2003a, 2003b, 2003c; Gruber 2006; Grek 2009). The OECD itself collated the press coverage of the first edition of PISA from the publication of the results in December 2001 until December 2002 (TiBi 2002), and the content has been summarised qualitatively by Gerlind Schmidt (2003b, 2003c). The numbers of articles varied greatly from

country to country: according to Karl Gruber (2006) the OECD website listed 8 pages of article titles from Finland (the highest ranked country), 88 pages from the UK, 32 from France, and 774 pages from Germany. Sotiria Grek (2009) compared the responses to PISA 2000 and 2003 in Finland, Germany and the UK and described the British press response to PISA 2006 (Grek 2008). Michelle Stack (2007) discussed US media coverage of PISA and TIMSS in terms of political discourse. However, ours is the first analysis that explicitly links the two themes of international rankings and negativity.

To explore these themes we undertook a systematic coding of press reports of the 2006 PISA and PIRLS editions from Britain, France and Germany and of the 2006 PISA edition from Finland. We interpret the results in the light of the different political, educational and journalistic systems in the four countries.

2. Analysis and Results

We tested our hypotheses by content analysis of a sample of articles from national and regional newspapers from France, Finland, Germany and Britain that were concerned with the PISA and PIRLS international educational rankings in 2007-08. The methodology used for sampling and coding the articles is described in the Appendix.

Results of the Press Article Analysis

Number of articles and sampling strategy

The number of relevant articles published in all four countries peaked in December 2007, the month of publication of the 2006 PISA study (Figure 1). A total of 912 German, 175 Finnish, 60 British, and 29 French articles were found from the search of six newspapers from each country for the year June 2007 to May 2008. Our initial aim was to compare 100 articles from each country, but Britain and France did not yield this number of articles in total, so all of the British and French articles were coded. Articles were sampled randomly from the German and Finnish search results (stratified by month), so that about 10 per cent of German articles from each month (100 articles in total, of which 5 were found by the coders to be irrelevant) and 57 per cent of Finnish articles (100 in total) were coded. The following discussion is based on this sample.

Figure 1 about here.

Figure 1. Number of relevant articles from our search of German, Finnish, British and French newspapers from 2007-08. See Appendix for details of search terms and newspapers.

The coded indicators

(a) Tenor

We use the term ‘tenor’ to mean the negativity or positivity of each article. A negative tenor denoted either a pessimistic view of the country’s performance (*‘Britian slumps in world league table’*) or criticism of the country’s education system (*‘la dégradation du système scolaire français’*, *‘die grundlegenden Probleme im Bildungsbereich’*) while a positive article expressed the results in a favourable light or emphasised positive aspects of the education system (*‘Suomen kannattaa kehittää koulutusosaamisesta vientituote’* (*‘Finland should use its excellent school system as a new export industry’*)). A neutral code reflected either a balance of views or a neutral, descriptive tone (*‘Suomi jälleen ykkönen Pisa-koulututkimuksessa’* (*‘Finland in the first*

position in PISA rankings’)). In Germany, Britain and France negativity towards each country’s education system and/or its performance in the PISA or PIRLS study significantly outweighed positivity ($p < 0.001$, binomial frequency distribution), while in Finland there were slightly more positive articles than negative ($p = ns$). Germany had the highest percentage of neutral articles (47 per cent) and Finland the least (6 per cent). Strikingly, the percentage of negative articles in the German and Finnish samples was the same (40 and 41 per cent of the articles respectively) despite their very different performances in the PISA study. British and French articles on average were much more negative than German or Finnish ones ($p < 0.001$, chi-squared test). Indeed apart from one British article in April 2008, no positive articles were found from Britain and France, compared to 13 per cent of German articles and 53 per cent of Finnish articles. The variation of the tenor with time is shown in Figure 2 and the coding results are shown in Table 2.

Table 2 and Figure 2 about here

Table 2: Results of content analysis. † $P < 0.001$, chi-squared test: null hypothesis is that the frequency distribution is identical between the countries.

Figure 2. Number of negative, neutral and positive articles in the coded samples by country and date.

(b) Criticism of the PISA and PIRLS studies

The articles were also scored on their attitude to the study itself, to investigate whether there was more criticism of the relevance or methodology of the study in countries that performed poorly. In fact we found relatively few articles in our sample expressing criticism of the PISA or PIRLS studies themselves, their methodology, their theoretical underpinning, or their ability to measure educational performance accurately in a wide variety of countries. If anything, the best performing country expressed the most criticism of the study (Finland, 20 per cent of articles) while Germany, Britain and France ranged from 7 to 12 per cent, though the difference was not significant (chi-squared test).

(c) Subject matter

Germany differed from the other countries in that 23 per cent of German articles were about topics other than school education, topics as diverse as television shows and the dangers of electromagnetic radiation, while all of the British or French articles were about school-level education, and only four per cent of Finnish articles were on other topics. It was clear from the articles that PISA and PIRLS were not familiar enough to their readers for British or French journalists to mention them in articles about unrelated subjects, in contrast to Germany, where PISA (though not, apparently, PIRLS) had entered the national consciousness. For instance, in a German article about architecture, the OECD is described as ‘*auch für die PISA-Studie verantwortlich*’ (‘*also responsible for the PISA study*’), that is, PISA was judged to be more familiar than the OECD to the readers. The proportion of German articles about other topics rose to 32 per cent in November and December 2007, so even in the months that the test results were published, the PISA study was mentioned in a wide range of articles and not just those that focused on school education or the rankings results.

(d) Coverage of PISA and PIRLS

In Germany, Britain and France, PISA and PIRLS were covered to different extents. PISA featured in all of the German articles coded. PIRLS (or IGLU, as it is known in Germany) was

mentioned in only 4 per cent of articles, and always mentioned together with PISA, whereas almost a third of British and French articles mentioned PIRLS, either alone or with PISA.

(e) Comparison with other countries

We also found differences in the degree to which other countries were compared with the ‘home’ country in the press articles. A large majority of Finnish and German articles mentioned no other countries (85 and 77 per cent respectively) while more than half of British and French articles mentioned the results of other nations to place the ‘home’ results in context.

(f) Political credit or blame

To test hypothesis (3), the articles were also coded as to whether politicians were mentioned, blamed, or credited for the test results, and for their responses to criticism, if any. Politicians were named in between 20 and 40 per cent of articles, depending on country. The level of blame was highest in Britain, politicians being criticised in over 60 per cent of the articles in which they appeared, and lowest in Finland, where none of the sample of 100 articles expressed personal criticism. In contrast, Finland had the highest level of credit of politicians, where 40 per cent of the articles that mentioned politicians did so approvingly.

Comparison of Germany 2001-02 and 2007-08

We compared 63 German articles from Dec 2001 – Jan 2002 from FAZ, TAZ and Die Welt with 57 articles from the same three newspapers in Dec 2007 – Jan 2008. We found that while negativity strongly outweighed positivity in both periods, the reaction was less negative in 2007-08 than in 2001-02 (chi-squared, $p < 0.001$). Criticism of the PISA methodology increased over this period in our sample (chi-squared, $p < 0.01$).

Discussion

Testing the hypotheses

Cross-national comparison of media, as with other social institutions, is an activity that presents many difficulties and challenges (Esser and Pfetsch 2004). We attempted in our selection of countries, newspapers and topics to assess a range of attitudes towards education and the printed media, and to choose newspapers that represented various political viewpoints and regional coverage. Given the very different media traditions in the four countries, a selection of only six newspapers from each country will lead to an imperfect comparison, and we recognize that the results could also depend on our exact choice of newspapers in each country. The newspapers within each country, however, had fairly similar distributions of negative, positive and neutral articles, with no newspaper appearing as a particular outlier within its country group. In an attempt to avoid coding bias due to differences in national cultural assumptions, the French, British and German newspaper articles were each coded by at least two coders with different native languages. The Finnish articles were all coded by two native Finnish speakers, who also coded some of the British articles to assess coder agreement. Any disagreements were resolved by discussion, resulting in the consensus coding used in the analysis. Given these caveats, we can draw only broad conclusions from this relatively limited number of countries, rankings, newspapers and articles.

So, to turn back to our first hypothesis:

(1) Negativity is found in press reports of educational rankings

We found that negativity significantly outweighed positivity in Germany, France and Britain. Even in Finland, whose PISA results were excellent, over forty per cent of the articles were judged to be negative, the same proportion as in Germany (though Finland had a higher proportion of positive articles). Thus we can confirm that press negativity is not just a British phenomenon as we found evidence of negativity in the press coverage of all four countries in this study. Hypothesis 1 is therefore clearly supported. The extent of this negativity, however, varied among the countries. The proportion of negative articles was significantly greater for France and Britain than for Germany and Finland. However, owing to the far higher total number of articles in Germany, the total number of negative articles was far higher in Germany than in the other countries (the continuing dependence of the results on socio-economic factors was a frequent theme). Finland, too, had a much larger number of articles (both negative and positive) than Britain or France, though not as high as Germany's total.

The negative press response to Germany's 2006 PISA rankings is striking, given the widespread interest in PISA in Germany. This salience dates back to the 2000 PISA results which revealed an unexpectedly poor performance among Germany's 15-year-olds compared to pupils in other OECD countries. Germany was also among the countries that had the highest dependence of educational results on socio-economic background (OECD 2001). These results contrasted starkly with Germany's own view of its education system and caused a widespread upheaval (the so-called '*PISA-schock*') (Helwig 2002; Gruber 2006), resulted in what Hubert Ertl termed 'a – by German standards – frantic reform agenda' (Ertl 2006). Although there had been little time to see fruits of any reforms, press interest was very high in the months preceding as well as following the release of the 2006 results (see Figure 1), and the salience of education as an issue was high (Table 3).

Given that we find support for our first hypothesis, we turn to our second hypothesis:
(2) *Negativity bias is found in press reports of educational rankings.*

Was this negativity simply due to objectively poor performance in the rankings, or was there evidence of *negativity bias*, that is, negativity that was unrelated to performance? It is of course well documented that press reports have a generally negative slant ('bad news sells papers', see for instance Cohen 1983; Soroka 2006) but it is also often observed that negativity can vary, for instance decreasing in 'honeymoon periods' as against other times.

To detect negativity bias (as opposed just to negativity) we would need to show that negativity was unrelated to performance across the range of countries or that even when performance was good or improving, negativity predominated. We tested Hypothesis 2 in three ways. First, we found that the Janis-Fadner coefficient of imbalance (a parameter that takes numbers of negative, positive and neutral articles into account (Janis and Fadner 1965)) was related to the performance of each country. That is, the worst performing countries (Britain and France) had the most negative coefficients, and Finland (the best performing country) had the most positive. This result, therefore, does not support Hypothesis 2.

But did any individual countries exhibit negativity bias? In our study, one country demonstrated unequivocally excellent performance. In PISA 2006, Finnish 15-year-olds repeated or exceeded

the excellent results of preceding cohorts in PISA 2000 and 2003 (Table 1), being placed first or second in all three subjects tested. In this case, we find some evidence for negativity bias. Although we found slightly more positive than negative Finnish articles, forty-one percent of the articles were coded negative, about the same percentage as in the German articles, a country ranked many places below Finland in the PISA results. Thus this result supports Hypothesis 2. Negativity bias towards the 2000 edition of PISA was similarly noted in some countries' coverage by Schmidt (2003c), who reported that even in highly ranked countries such as Canada, Sweden and the Netherlands, self-critical reactions dominated the response to this study. Issues that were criticised in our sample of the Finnish press included the continuing difference in the attainment of boys and girls; the standardisation of education leading to the neglect of 'excellence'; the lack of comparably good results of Finnish higher education; and the concentration on a narrow range of knowledge in the school curriculum, with insufficient attention paid to social interaction and practical skills, and the neglect of indicators of social alienation. This last issue may have been particularly salient in Finland at that time as the country had experienced a school shooting tragedy in November 2007 (Raittila et al. 2010).

Germany provided a third, though imperfect, test of negativity bias. Germany's absolute scores and positions had improved in 2006 compared to 2000 and 2003. However, this improvement was far from clear-cut as the socio-economic dependence of the results had not improved, and the OECD PISA report (OECD 2007) pointed out that Germany's overall improvement was not statistically significant. This lack of statistical significance was, however, rarely mentioned in the press articles in our sample which took the improvements largely at face value. We undertook an analysis of articles from three newspapers in 2001-02 with the same newspapers in 2007-08. We found that the articles were significantly less negative in 2007-8, and we also found that the level of criticism of the PISA study increased. Examples of such criticism can be found in the German press (e.g. Kahl 2007) and in the academic literature (e.g. Prais 2007). Thus this comparison gave no support for negativity bias, since the coverage became less negative as the results (somewhat) improved.

So from our analysis, we found that negativity was universal in the press of these four countries (Hypothesis 1 was supported), and that there was (arguably) some evidence of negativity bias (Hypothesis 2 was supported in one of the three tests that we applied). We found little evidence of the PISA or PIRLS studies themselves being held to be irrelevant or flawed. This led us to ask what impact the negativity of the press articles might have on politicians. Was the negativity 'anonymous', that is, directed towards 'the education system' or 'the results', or were named politicians held responsible?

We tested this question with our third hypothesis:

(3) Politicians are personally blamed in the press reports of educational rankings

We found that politicians were treated very differently by the press of the four countries in their coverage of these international surveys, based on our sample. These differences no doubt reflect the different political systems and the extent to which central politicians are held responsible for education policy, as outlined below, but may also reflect the different press traditions of the countries.

While it is a stereotype to characterise Anglo-Saxon and German journalists as ‘bloodhounds’ and ‘missionaries’ respectively (Köcher 1986), French journalists as ‘literary philosophers’ (Benson 2002; Palmer 2008), and Finnish journalists as ‘watchdogs with a conscience’ (Heinonen 1997), and despite some convergence of European media attitudes (Esser 1999; Hallin and Mancini 2004a, Marchetti 2009) there are still differences between the journalistic traditions of these countries (Weaver 1997; Esser 1998; Hallin and Mancini 2004b; Palmer 2008; Curran et al. 2009). The British press is widely portrayed as being more combative and investigative, and the German, Finnish and French press as more analytical and less critical, with more of a role to explain complicated issues to the public than to challenge political policies. There is also traditionally less deference shown to interviewees in Britain. The lack of deference is epitomised by the British tabloid press which ‘regards elected officials mostly as scoundrels, liars or fools’ (Rubin 2007).

Not only do press traditions vary between the four countries, but the levels of political control of the education systems vary as well. As noted above, our coverage of each country’s press was necessarily incomplete, but we still found evidence in our sample of articles that there were differences in the ways in which politicians were held responsible for educational success. To place our findings in context, we summarise some features of the political and institutional aspects of the education systems of each country in Table 3.

Table 3 about here

Table 3. Characteristics of the Countries’ School Systems in 2007.

The degree of blame found in the press articles is clearly influenced by a range of factors including the rankings results themselves, the tenor of the media towards those results, the attitudes and traditions of that country’s press, and the degree to which politicians are held responsible for educational performance. We cannot attempt to disentangle these factors for just four countries, but suggest that this would be a fruitful area for investigation. We offer some observations, however, which may shed light on these particular cases and highlight differences between the countries.

Britain had the most personal blame in the press articles and is also the country with the least deferential press, and the one where government education ministers (particularly under Tony Blair’s Labour government) took personal responsibility for educational performance (Barber 2007). In England (and to a lesser extent in the other UK countries) central government controls many aspects of education, imposing directives, national testing regimes, targets and league tables on schools (despite the fact that local authorities are nominally responsible for delivering education to children in their area). There were clear lines of responsibility to the Westminster education minister (Ed Balls in 2007) and to the Scottish and Welsh schools ministers. We note, however, that the press interest, while highly negative, was relatively short-lived, and there were no calls for ministerial resignations.

France has the most centralised education system among our cases., However, while the minister for education (Xavier Darcos) was frequently mentioned in press articles (41 per cent), he was rarely personally criticized for details of his education policies. The highly negative (though few in number) press articles criticised ‘the school system’ and bemoaned the poor

results but did not personalize this blame. This may reflect the more deferential attitude of the press in France.

The success of Finland in PISA 2000 (which came as a surprise even to educationalists at the time (Grek 2009; Välijärvi et al. 2002)) was repeated in 2006. In our press articles we found over 20 different politicians mentioned in the context of the PISA rankings. The education minister at the time, Sari Sarkomaa, was mentioned in only 6 per cent of articles, of which two-thirds were favourable to her, and the rest were neutral. Alone among the countries in our study, Finnish politicians were credited far more than they were blamed. Even the negative articles did not tend to blame politicians, which may reflect politico-institutional features of the Finnish education system. That is, local government is responsible for providing school education, and local responsibility is emphasised by the Finnish political system. This has not led, however, to a diversity of educational policies, but to a fairly standardized school system, regulated and steered by central government by means of many direct and indirect instruments. The pathways for assigning credit and blame are therefore complicated and obscure. National politicians are rarely credited because local governments are in charge of the actual schooling. On the other hand, blame is also difficult to assign because individual schools (or municipalities) are not identified in the PISA scores and it may therefore make sense to criticize 'the system' when lacking more specific targets.

In Germany, responsibility for education is devolved almost completely to the *Länder*. Each *Land* has responsibility for curricula, teacher training, school inspection and examinations and has its own education minister. The Conference of Education Ministers (*Kultusministerkonferenz*, KMK) was heavily involved in proposing reforms after the first round of PISA in 2000 (Klieme and Döbert 2003; Klieme et al. 2004; Sroka and Döbert 2004; Ertl 2006). Reflecting this lack of involvement of central government in the area of school education, the Federal Education Minister Annette Schavan was mentioned in only 3 of the 95 articles in our sample (and none in the immediate aftermath of the PISA results) and was in no case held responsible for educational outcomes. Regional education ministers were mentioned far more often (in about 20 per cent of articles) and were afforded both praise and blame for their educational policies. It should be noted, however, it was impracticable and outside the scope of this paper to sample the extensive German regional press to include newspapers from every *Land*. A more comprehensive survey would be needed to determine the overall press attitude to regional politicians and their education policies.

Therefore we find most support for Hypothesis 3 in Britain, where politicians were blamed in 23 per cent of the articles. The reasons for the relative lack of blame in the other countries is very likely due to a combination of factors, discussed above. But even in the absence of personal blame, negative press coverage may elicit a political response and it is to that question that we turn with our final hypothesis:

(4) Educational rankings produce a political response that is influenced by the intensity and tenor of the press coverage

Rankings have the (implicit or explicit) aim of driving or assisting performance improvement, and this aim is apparent in many of the OECD's and the IEA's publications. It is well

documented that rankings can lead to radical changes in behavior of the ranked organizations in some contexts. For instance, Sauder and Espeland (2009) showed that US law schools, even if initially sceptical, found that the USN rankings became so important that they had to devote increasingly large resources both to monitoring the relevant indicators and seeking to improve them. Chatterji and Toffel (2010) found that companies ranked low on environmental ratings made more effort to improve than companies ranked higher. In neither case did low-ranked institutions simply ignore the rankings. Did PISA and PIRLS have similar effects?

We should here distinguish between the immediate political response that we measured in our study, and long-term performance improvement that is beyond the scope of this paper. Our data set allows us to distinguish between the rankings results themselves and the way the media presented those results to the public and politicians. Given that the results were presented differently in the four countries, what political responses did the rankings elicit?

As discussed earlier, these international educational rankings can only exert their effects through governments, since not all schools are tested and the results of individual schools are not revealed. But how strongly might we expect this mechanism to operate? If the PISA results are poor there are limited opportunities for political credit-claiming. Blame-avoidance also seems to be a weak mechanism, given that ministerial turnover is generally rapid (compared, for instance, to university deans or company chief executives) and ministers can generally blame a previous administration, or the local delivery of services, or promise reforms in the knowledge that the next survey will not be for another three years. And the pressure that ministers can apply to schools based on these rankings is likely to be a weaker mechanism for improvement than more direct incentives such as school league tables or targets on which head-teachers' jobs depend. So we might expect that the PISA and PIRLS *results* would have a relatively weak direct effect, and actually not trouble politicians too much. But the *way in which those results are presented by the media* might exacerbate or reduce their effect in eliciting a response (as has been shown by Soroka (2006) in the case of economic information).

With just four cases we cannot of course quantify these effects. But qualitatively, we can classify the political responses that we detected in the press articles as minimal (France and Finland), presentational (Britain) and reform-orientated (Germany). Even this tiny sample shows that there is no simple relationship between results and responses, since France with the 'worst' results and Finland with the 'best' had similarly weak responses.

Considering the responses in the light of the way in which the results were presented illuminates these issues. Finland had a high level of press interest in its continuing excellent results and a surprisingly high level of negativity (though this was not personalized, and was outweighed by the number of positive articles), but no specific responses by politicians. This lack of political response is of course not surprising given the good results and absence of blame from the press, and would be a rational response by Finnish politicians. We found a relatively low level of media interest in Britain and France though such as there was, was highly negative. Interest peaked around the time of the release of results but fell off quickly thereafter. We found a high level of blame of individual politicians in the British press, and many individual responses by politicians to that blame. The responses were mainly presentational (Hood 2010), such as deflecting blame to previous administrations, teachers and even parents, rather than announcing specific policy

changes. In France the political response was essentially undetectable (the minister was quoted as ‘expressing his disappointment’ in the results but not taking responsibility for them). In Germany in 2007-8, as at the time of the PISA 2000 results (Gruber 2006), there was overwhelming and sustained press interest, and we found both credit and blame of politicians. Many articles discussed the various reforms initiated in response to earlier editions of PISA, and often criticized their apparent lack of success. The political responses were to defend existing reforms or to suggest new ones.

We therefore find some qualitative support for our Hypothesis 4, though of course far more cases would be needed to generalize these findings. That is, the political response was muted in France and Finland where the press coverage was either sparse (France) or reasonably positive (Finland). In Britain negative press coverage that personally criticized ministers was met by blame-avoidance strategies. In Germany where the press response was vast in number (and 40 per cent of those articles were negative), politicians were more likely to embrace policies for reform. In these cases, therefore, we suggest that the political response was influenced as much by the intensity and tenor of media interest as by the results themselves. So we conclude that while such international rankings may not be strong levers for reform, the strength of such levers does appear to depend on the way the results are presented. Clearly, investigating interactions between the results and the tenor of the press coverage further, and accounting for differences in political, institutional and press factors, would require a far larger number of countries than is presented here, but such an investigation would be a promising avenue for future research.

So why did Germany put a huge political effort (post-2000, and continuing today) into responding to what we argue is a relatively weak incentive? We can’t answer that directly from our data set, but suggest that the PISA results fed into various debates about social mobility and disparities between *Länder*, and in the ongoing educational debates about for instance, selective versus comprehensive education, whole-day schooling, and teacher training methods (Kotthoff 2011). As we discussed in the introduction, the strength of public opinion as expressed in the media is likely to influence political responses. The overwhelming media response in 2001 is cited by many (e.g. Gruber 2006; Grek 2009; Rotte and Rotte 2007; Ertl 2006) as the reason for the political response at that time, and our study shows that the high level of press interest continued at least until 2008. Education was also a much more salient issue for the public in Germany at that time than in the other countries in our study, as shown in Table 3. The strength of the media response may also relate to memories of Germany’s *Bildungskatastrophe* in the 1960s. This term was coined by George Picht (1964) who argued that the education system was underfunded, unsuited to Germany’s industrial economy and in need of radical reform. Debate about education policy has continued since that time, and although it led to reforms such as the foundation of comprehensive schools in some *Länder* in the 1970s, it did not fundamentally change Germany’s selective tripartite school system that still caters for the majority of pupils (Wilde 2002). The ‘*PISA-schock*’ in 2001 gave ammunition to various groups to push their own agenda for reform, but as we have seen from our analysis of articles from 2007 and 2008 no widespread consensus has developed as to what the ideal education system would be. This may be why despite a variety of initiatives attributed to PISA since 2000, there has been no single response to PISA across Germany (Kotthoff 2011; Rotte and Rotte 2007). A study of political communication effects (Schmitt-Beck 2004) showed that exposure to political reporting in the media had a greater effect on voting intention in Germany than in Britain or the United States in

the early 1990s. That is, newspapers and television had a strong influence on public opinion in Germany. If that was still true in the 2000s we can argue that the intensity and negativity of the German press response to PISA, together with the salience of education as an issue, was likely to provoke a strong political response, as indeed we found.

Conclusion

In our cross-national analysis of the press response to educational rankings in four EU countries, we found that negativity was indeed found in press reports of such rankings in all four countries (our hypothesis 1). This supports the findings of other studies of policy areas such as government targets in Britain (Hood and Dixon 2010, James 2004) and economic performance in the US (Soroka 2006), and shows that negativity is not just an Anglo-Saxon phenomenon, and occurs even when performance was good. We found less clear results for negativity bias (hypothesis 2), and would need more cases for a satisfactory quantitative test of negativity bias, but our method could be applied to future rounds of international rankings.

While there was evidence for negativity in press reports in all four countries, there was relatively little blame directed personally at politicians (hypothesis 3). The exception to this was Britain and specifically England, where politicians did appear to be held responsible for the performance results. In terms of political response, we found that ministers in Britain showed the most direct response to criticism (but with largely presentational strategies rather than substantial educational reforms) and in Germany the response was to discuss educational reforms. We argue that, at least in part, the way in which the results were presented in the media (the intensity of press coverage, and the negativity of the articles) influenced the political responses that we observed (hypothesis 4).

Considering our results in the context of the literature that predicts that rankings will lead to reactions by those ranked, we see that Germany undertook a radical educational reform agenda in response to PISA 2000 which continued through the period covered by our study, while the relatively worse performance of France, for instance, did not lead to a similar response. We therefore find that PISA and PIRLS do not fit neatly into the conclusion from studies of other kinds of rankings (that rankings tend to elicit responses from those ranked), and we have to invoke other factors than performance such as political and institutional factors and the intensity and tenor of the media response to suggest explanations for the different political responses. As we have argued, we would expect such international rankings to exert a relatively weaker effect on either school performance or educational policy than, for instance, domestic or international rankings of universities do on those institutions. The value of such rankings may be more as a form of 'intelligence' (Hood 2007) that allows educationalists to better understand characteristics of successful school systems. Thus the role of such rankings as a 'lever for improvement' is likely to be indirect.

In conclusion, our study has shown that it is possible to analyse a cross-national set of press articles in order to investigate the variety of responses to international rankings. We have shown that there is no simple relationship between the results of such rankings and the political responses to those results, and we have highlighted the role of the media in influencing the political response. The effect of such rankings on educational policy and ultimately performance is a fertile area for research. Our study shows that the importance of the press in eliciting and

criticizing reform proposals should not be underestimated. On the contrary, further studies are needed to understand the interactions between the rankings results, the attitude of the media, the political responses and the eventual outcome of the rankings exercise in terms of educational performance.

References

- Arndt, C. 2008. 'The Politics of Governance Ratings', *International Public Management Journal*, 11, 3, 275-297.
- Barber, M. 2007. *Instruction to Deliver: Tony Blair, Public Services and the Challenge of Achieving Targets*. London: Methuen.
- Baumeister, R.F., E. Bratslavsky, C. Finkenauer and K.D. Vohs. 2001. 'Bad is Stronger than Good', *Review of General Psychology*, 5, 4, 323-370.
- Benson, R. 2002. 'The Political/Literary Model of French Journalism: Change and Continuity in Immigration News Coverage, 1973–1991', *Journal of European Area Studies*, 10, 1, 49-70.
- Besançon, M. 2003. *Good Governance Rankings: The Art of Measurement*. WPF Reports No. 36. Cambridge: World Peace Foundation.
- Boyne, G.A., O. James, P. John and N. Petrovsky. 2010. 'Does Public Service Performance Affect Top Management Team Turnover?' *Journal of Public Administration Research and Theory*, 20, Suppl. 2, i261-i279.
- Bradshaw, J., L. Sturman, H. Vappula, R. Ager, and R. Wheater. 2007. *Achievement of 15-year-olds in England: PISA 2006 National Report*. Slough: NFER.
- Chatterji, A.K. and M.W. Toffel. 2010. 'How Firms Respond to Being Rated.' *Strategic Management Journal*, 31: 917-945.
- Cohen, B. 1983. 'Nuclear Journalism: Lies, Damned Lies, and News Reports', *Policy Review*, 26, 3, 70-74.
- Curran, J., I. Salovaara-Moring, S. Coen and S. Iyengar 2009 'Crime, foreigners and hard news: A cross-national comparison of reporting and public perception' *Journalism* 1,11, 1–17.
- DfES (Department for Education and Skills). 2004. 'Programme for International Student Assessment (PISA) 2003: England Sample and Data'. SFR47/2004. London: DfES.
- Ertl, H. 2006. 'Educational Standards and the Changing Discourse on Education: the Reception and Consequences of the PISA Study in Germany', *Oxford Review of Education*, 32, 5, 619-634.

Espeland, W.N. and M. Sauder 2007. 'Rankings and Reactivity: How Public Measures Recreate Social Worlds', *American Journal of Sociology*, 113, 1, 1-40.

Esser, F. 1998. 'Editorial Structures and Work Principles in British and German Newsrooms', *European Journal of Communication*, 13, 375-405.

Esser, F. 1999. "'Tabloidization' of News: A Comparative Analysis of Anglo-American and German Press Journalism", *European Journal of Communication*, 14, 291-324.

Esser, F. and B. Pfetsch (Eds). 2004. *Comparing political communication: theories, cases, and challenges*. Cambridge: Cambridge University Press.

Frederickson, H.G. and E.C. Stazyk. 2010. 'Ranking US Public Affairs Educational Programmes' in H. Margetts, P. 6 and C. Hood (eds) *Paradoxes of Modernization: Unintended Consequences of Public Policy Reform*. Oxford, Oxford University Press, pp. 63-80.

Grek, S. 2009. 'Governing by Numbers: the "PISA Effect" in Europe', *Journal of Education Policy*, 24, 1, 23-37.

Grek, S. 2008. 'PISA in the British Media: Leaning Tower or Robust Testing Tool?' CES Briefing No. 45.

Gruber, K.H. 2006. 'The German "PISA-shock": Some Aspects of the Extraordinary Impact of the OECD's PISA Study on the German Education System', in H. Ertl (ed.) *Cross-national Attraction in Education: Accounts from Britain and Germany*. Oxford, Symposium Books, pp. 195-208.

Hallin, D.C. and P. Mancini. 2004a. 'Americanization, Globalization and Secularization: Understanding the Convergence of Media Systems and Political Communication', in F. Esser and B. Pfetsch (eds) *Comparing Political Communication: Theories, Cases, and Challenges*. Cambridge: Cambridge University Press, pp. 25-44.

Hallin, D.C. and P. Mancini. 2004b. *Comparing Media Systems: Three Models of Media and Politics*. Cambridge: Cambridge University Press.

Harrison, S. and C. Pollitt (eds). 1992. *Handbook of Public Services Management*. Oxford: Blackwell.

Heinonen, A. 1997 'The Finnish journalist: Watchdog with a Conscience' in D. Weaver (ed) *The Global Journalist*, Cresskill: Hampton Press.

Helwig, G. 2002. 'PISA-Schock', *Deutschland-Archiv*, 35, 3, 369-372.

Hood, C. 2007. 'Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles?' *Public Money and Management*, 27, 2, 95-102.

Hood, C. 2010. *The Blame Game: Spin, Bureacracy and Self-Preservation in Government*. Princeton and Oxford: Princeton University Press.

Hood, C., R. Dixon and C. Beeston. 2008. 'Rating the Rankings: Assessing International Rankings of Public Services', *International Public Management Journal*, 11, 3, 298-328.

Hood, C, W. Jennings, R. Dixon, B. Hogwood and C. Beeston. 2009. 'Testing Times: exploring staged responses and the impact of blame management strategies in two exam fiasco cases'. *European Journal of Political Research*, 48,6, 695-722.

Hood, C. and R. Dixon. 2010. 'The Political Payoff from Performance Target Systems: No-Brainer or No-gainer?' *Journal of Public Administration Research and Theory*, 20, Suppl. 2, i281-i298.

Hutchison, D. and I. Schagen. 2007. 'Comparisons between PISA and TIMSS - Are We the Man with Two Watches?' in T. Loveless (ed.) *Lessons Learned: What International Assessments Tell Us about Math Achievement*. Washington: Brookings Institution Press.

Janis, I.L. and R. Fadner. 1965. 'The Coefficient of Imbalance', in H. Lasswell (ed.) *The Language of Politics*. Boston: MIT Press, pp. 153-69.

James O. 2004. 'The UK Core Executive's Use of Public Service Agreements as a Tool of Governance', *Public Administration*, 82, 2, 397-419.

James O. 2011. 'Managing Citizens' Expectations of Public Service Performance: Evidence from Observation and Experimentation in Local Government', *Public Administration*, 89, 4, 1419-1435.

James, O. and P. John. 2007. 'Public Management Performance Information and Electoral Support for Incumbent English Local Governments', *Journal of Public Administration Research and Theory*, 17, 4, 567-580.

Kahl, R. 2007. 'Leistungen Mittelmaß, Chancengleichheit schlecht', *Die Zeit*, 49 (29.11.2007); 'Vorschneller Datenerguss', *Die Zeit*, 49 (29.11.2007); 'Das Ringen um PISA', *Die Zeit*, 51 (13.12.2007).

Klieme, E., H. Avenarius, W. Blum and 8 others. 2004. *The Development of National Educational Standards. An Expertise*. Berlin: BMBF (Bundesministerium für Bildung und Forschung)

Klieme, E. and H. Döbert. 2003. *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten*. Berlin: BMBF (Bundesministerium für Bildung und Forschung) (published in English as Sroka and Döbert 2004)

Kotthoff, H.-G. 2011. 'Between Excellence and Equity: the Case of the German Education System'. *Revista Española de Educación Comparada*, 18, 27-60.

- Köcher, R. 1986. 'Bloodhounds or Missionaries: Role Definitions of German and British Journalists', *European Journal of Communication* 1, 1, 43-64.
- Krippendorff, K.H. 1980. *Content Analysis: An Introduction to Its Methodology* (and 2nd ed. 2004) Beverley Hills: Sage.
- Lacy, S. and D. Riffe. 1996. 'Sampling Error and Selecting Intercoder Reliability Samples for Nominal Content Categories', *Journalism and Mass Communications Quarterly*, 73, 4, 963-973.
- Lau, R.R. 1982. 'Negativity in Political Perception', *Political Behaviour*, 4, 4, 353-377.
- Marchetti, D. 2009. 'The revelations of investigative journalism in France', *Global Media and Communication*, 5, 368-388.
- Mullis, I.V.S., M.O. Martin, A.M. Kennedy and P.Foy. 2007. *PIRLS 2006 International Report: IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries*. Chestnut Hill: IEA.
- OECD (Organisation for Economic Co-operation and Development). 2001. *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD.
- OECD (Organisation for Economic Co-operation and Development). 2004. *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.
- OECD (Organisation for Economic Co-operation and Development). 2007. *PISA 2006: Science Competencies for Tomorrow's World. Volume 1 – Analysis*. Paris: OECD.
- Palmer, M. 2008. 'International News from Paris- and London-based Newsrooms', *Journalism Studies*, 9, 5, 813-821.
- Peter, J. and E. Lauf. 2002. 'Reliability in Cross-National Content Analysis', *Journalism and Mass Communication Quarterly*, 79, 4, 815-832.
- Prais, S. J. 2007. 'Two Recent (2003) International Surveys of Schooling Attainments in Mathematics: England's Problems', *Oxford Review of Education*, 33, 1, 33-46.
- Picht, G. 1964. *Die deutsche Bildungskatastrophe*. Olten: Walter.
- Raittila, P., K. Koljonen and J. Väliverronen. 2010 *Journalism and school shootings in Finland 2007-2008*. Tampere: University of Tampere Press. <http://tampub.uta.fi/tup/978-951-44-8257-1.pdf>
- Riffe, D. and A. Freitag. 1997. 'A Content Analysis of Content Analyses: Twenty-five Years of Journalism Quarterly', *Journalism and Mass Communications Quarterly*, 74, 4, 873-882.

- Rotberg, R.I. and D.L. West. 2004. *The Good Governance Problem: Doing Something About It*. WPF Reports No. 39. Cambridge: World Peace Foundation.
- Rotte, R. and U. Rotte, U. 2007. 'Recent Education Policy and School Reform in Bavaria: A Critical Overview', *German Politics* 16, 2, 292-313.
- Rozin, P. and E.B. Royzman. 2001. 'Negativity Bias, Negativity Dominance, and Contagion', *Personality and Social Psychology Review*, 5, 4, 296-320.
- Rubin J.P. 2007. 'Backstage at Downing Street', *New York Times* 2 Sept 2007.
- Sauder, M. 2008. 'Interlopers and Field Change: The Entry of U.S. News into the Field of Legal Education', *Administrative Science Quarterly* 53, 209-234.
- Sauder, M., and W.N. Espeland. 2009. 'The Discipline of Rankings: Tight Coupling and Organizational Change', *American Sociological Review* 74, 63-82.
- Schmidt, G. 2003a. 'Reaktionen auf die PISA-Ergebnisse in anderen europäischen Staaten: Eine Analyse von Presseberichten', *päd forum*, 31, 3, 159-164.
- Schmidt, G. 2003b. 'PISA – Reaktionen in teilnehmenden Ländern im Spiegel ihrer Press: Vergleich', *Trends in Bildung international (TiBi)*, 5. 1-7.
- Schmidt, G. 2003c. 'Reactions of Participating Countries as Reflected in their Press', *European Education*, 35, 4, 58-69.
- Schmitt-Beck, R. 2004. 'Political Communication Effects: The Impact of Mass Media and Personal Conversations on Voting' in F. Esser and B. Pfetsch (eds.), *Comparing Political Communication, Theories, Cases, and Challenges*. Cambridge, Cambridge University Press, 293-322.
- Soroka, S.N. 2006. 'Good News and Bad News: Asymmetric Responses to Economic Information', *Journal of Politics*, 68, 2, 372-385.
- Sroka, W. and H. Döbert. 2004. *Features of Successful School Systems: A Comparison of Schooling in Six Countries*. Münster: Waxmann Verlag.
- Stack, M. 2007. 'Representing School Success and Failure: Media Coverage of International Tests', *Policy Futures in Education*, 5, 1, 100-110.
- TiBi (Trends in Bildung international). 2002. *PISA - Reaktionen der teilnehmenden Länder TiBi Nr.5*. The collated articles are no longer available on the OECD website (personal communication, OECD).
- Väljjarvi, J., P. Linnakylä, P. Kupari, P. Reinikainen, and I. Arffman. 2002. *The Finnish success in PISA – And some reasons behind it*. Jyväskylä: Institute for Educational

Research, University of Jyväskylä.

Van de Walle, S. 2006. 'The State of the World's Bureaucracies', *Journal of Comparative Policy Analysis*, 8, 4, 439-450.

Weaver, D. (Ed.) 1997. *The Global Journalist: News People Around the World*. Cresskill: Hampton Press.

Weaver, R.K. 1986. 'The Politics of Blame Avoidance', *Journal of Public Policy*, 6, 4, 371-98.

Wilde, S. 2002. 'Secondary Education in Germany 1990-2000: "One decade of non-reform in unified German education"?' *Oxford Review of Education* 28, 1, 39-52.

Appendix Methodology

Content analysis

Methodology for the content analysis reported here was guided by references such as Krippendorff (1980, 2004), Lacy and Riffe (1996), and Riffe and Freitag (1997) concerning hypothesis formulation, sample sizes, randomization of sources, selection of coding criteria and evaluation of coder reliability. The discussions of cross-national comparisons in Esser and Pfetsch (2004, particularly the chapter by Kleinstuber) and Peter and Lauf (2002) also informed our analysis.

Newspapers

The newspapers were chosen to include a variety of political orientations, regional and national coverage, and a variety of social demographics, together with high circulation figures. Most of the French, German and British papers were available on Lexis-Nexis, apart from Bild and Süddeutsche Zeitung which were searched using the newspapers' own archives. The Finnish newspapers were also searched from their own archives. Details of the newspapers are given in Table 4, together with the numbers of articles found. Due to access limitations, the searches of 2001-2002 German papers were of FAZ, Taz and Welt.

Table 4 about here

Table 4. Newspapers included in the 2007-8 search with their 2008 circulation figures and the number of articles found.

Dates searched

The PISA 2000 results were published in December 2001, PISA 2006 results in December 2007, and PIRLS 2006 results in November 2007. The archives were searched for the six months before and six months after these dates (i.e. 1 June 2001 to 30 May 2002 and 1 June 2007 to 30 May 2008).

Search terms

Germany: [PISA OR PIRLS OR IGLU OR Schul Leistungs Untersuchung OR Grundschul Lese Untersuchung]; Britain: [PISA OR PIRLS OR International Student Assessment OR Reading Literacy Study]; France: [PISA OR PIRLS OR Programme international pour le suivi des acquis

des élèves OR Capacités lecture des jeunes OR Programme international pour le suivi des acquis OR lecture jeunes]; Finland: [PISA].

Sampling

Articles judged to be ‘probably relevant’ were extracted from the total list of hits by individual assessment of the keywords and titles of each document. If the total number of probably relevant articles for a country was greater than 100 (Germany and Finland) a stratified sampling scheme was applied in which a percentage of each month’s ‘hits’ was randomly selected. The coded articles and the coding results are available from the authors on request.

Coding criteria

First, the actual relevance of the article was assessed, i.e. whether the article referred to one of the educational studies, was not a duplicate of another coded article, and was a newspaper article (rather than, for instance, a list of forthcoming articles or a figure caption). If it was judged irrelevant, the article was omitted from further consideration (this applied to about 5 per cent of the ‘probably relevant’ articles);

The following indicators were coded for all relevant articles:

- a. General tenor of the article towards the country’s results or its national (school-level) education system, on a three point scale (negative, neutral or positive); neutral included articles where negative and positive aspects were balanced, or the tone was neutral throughout;
- b. Criticism of the study (methodology, relevance, execution) on a two point scale (critical: –1; uncritical or supportive: 0);
- c. Subject matter of the article (whether the article reported the results only; mentioned or discussed the national school education system (including individual schools, regions or initiatives, but excluding university education, libraries, and private educational initiatives); or was wholly about other topics);
- d. Whether PISA, PIRLS or both studies were mentioned.
- e. Whether the article compared the ‘home’ country to other named countries, and whether those named countries performed better or worse than the ‘home’ country.
- f. Which articles mentioned national or regional politicians, what credit or blame was accorded to politicians, and any responses made by the politicians.

Basic information on each article was collected: date of publication, the name of the newspaper, ‘home’ country of the newspaper, author, title, and the length of the article in words. As different languages have different ‘efficiencies’ i.e. equivalent information is expressed in a different number of words, we did not draw any conclusions from this indicator, which in any case did not show significant differences between countries.

Agreement between coders

Initial coding criteria were agreed and ten articles were coded by four authors (RD, MM, CA and CH). The criteria were refined on the basis of this exercise, and twenty more articles from each of France, Britain and Germany were coded by RD and MM. Krippendorff’s alpha (Krippendorff 1980) was calculated for each indicator to assess coder agreement. Where an indicator had an alpha (agreement above chance) below 80 % (indicators ii, iv and v), the coding criteria were discussed and amended before the rest of the articles were coded. The descriptions

above reflect the final coding criteria. JV and KE-P coded the Finnish articles using the same criteria. They also coded twenty British articles for calculation of coder agreement with the other coders. They found the same percentage of negative and neutral articles as did RD and MM, showing no systematic bias, although there was a disagreement on 6 articles as to the classification, showing that there is a degree of uncertainty in the results. Every article was coded by at least two authors and, where necessary, a consensus coding was agreed for each article by discussion. The exception to this was that indicators of political blame and response were coded by one coder for each article.

Data analysis

The distribution of code frequencies between countries from the 2007-8 articles was compared using Pearson's chi-squared test (SPSS version 17). The null hypothesis was that each country had the same distribution of codes for a particular indicator. The 'general tenor' (indicator a) of the article was additionally analysed as follow: within countries, negativity was assessed by the binomial frequency distribution, that is, the probability that an unbiased sample will give this degree of imbalance of positive to negative codes by chance, ignoring neutral codes (Microsoft Excel 2003). The Janis-Fadner coefficient of imbalance (Janis and Fadner 1965) was also calculated. This coefficient takes into account the proportion of neutral codes and takes a value of -1 to +1.

Table 1. Performance in PISA and PIRLS. PISA tested 15-year-old school pupils in science, reading and mathematics, PIRLS tested 10-year-old pupils in reading (data from OECD 2001; OECD 2004; OECD 2007; Mullis et al. 2007).

Study	Year of test	Number of countries taking part	<u>Germany</u>	<u>Finland</u>	<u>United Kingdom</u> ^a	<u>France</u>
			Mark (position) [OECD average mark = 500]			
PISA - Science						
	2000	43	487 (20)	538 (3)	532 (4)	500 (12)
	2003	41	502 (15)	548 (1)	518 (13) ^b	511 (10)
	2006	57	516 (8)	563 (1)	515 (9)	495 (19)
PISA - Reading						
	2000	43	484 (21)	546 (1)	523 (7)	505 (13)
	2003	41	491 (18)	543 (1)	507 (9) ^b	496 (14)
	2006	57	495 (14)	547 (2)	495 (13)	488 (17)
PISA - Maths						
	2000	43	490 (19)	536 (4)	529 (8)	517 (10)
	2003	41	503 (16)	544 (1)	508 (15) ^b	511 (13)
	2006	57	504 (14)	548 (1)	495 (18)	496 (17)
PIRLS – Reading Literacy						
			<u>Germany</u>	<u>England</u> ^a	<u>Scotland</u> ^a	<u>France</u>
	2001	35	539 (11)	553 (3)	528 (14)	525 (18)
	2006	41	548 (11)	539 (19)	527 (26)	522 (27)

Notes:

^a All four countries of the United Kingdom (England, Wales, Scotland and Northern Ireland) took part in PISA and the mean results for the whole country were given in the PISA reports, although the individual country's scores were reported by the Department for Children, Schools and Families (Bradshaw et al. 2007). England and Scotland (but not Wales or Northern Ireland) took part in PIRLS, and the results were reported separately.

In the text of the paper, we use the term 'Britain' to include England, Scotland and Wales, as our search of papers did not include any from Northern Ireland (which is part of the United Kingdom of Great Britain and Northern Ireland).

^b The United Kingdom (specifically England) narrowly failed to meet the participation criteria for PISA 2003 and so was not included in the international tables for that year. The results were reported in the UK national report (DfES 2004) and are shown in this table for comparison.

Table 2: Results of content analysis. † $P < 0.001$, chi-squared test: null hypothesis is that the frequency distribution is identical between the countries.

	<u>Germany</u>	<u>Finland</u>	<u>Britain</u>	<u>France</u>
	Articles from June 2007 to May 2008			
	number of articles (percentage of articles)			
Tenor†				
Negative	38 (40 %)	41 (41 %)	47 (78 %)	20 (69 %)
Neutral	45 (47 %)	6 (6 %)	12 (20 %)	9 (31 %)
Positive	12 (13 %)	53 (53 %)	1 (2%)	0 (0 %)
Janis-Fadner coefficient of imbalance	-0.11	0.06	-0.60	-0.48
Subject matter†				
Results only	2 (2 %)	37 (37 %)	3 (5 %)	4 (14 %)
School education	71 (75 %)	59 (59 %)	57 (95 %)	25 (86 %)
Other topics	22 (23 %)	4 (4 %)	0 (0 %)	0 (0 %)
Which study? †				
PISA only	91 (96 %)	n/a	40 (67 %)	21 (72 %)
both	4 (4 %)		3 (5 %)	1 (3 %)
PIRLS only	0 (0 %)		17 (28 %)	7 (24 %)
Comparison with other countries†				
Yes	14 (15 %)	23 (23 %)	34 (57 %)	18 (62 %)
No	81 (85 %)	77 (77 %)	26 (43 %)	11 (38 %)
Criticism of study (ns)				
Yes	11 (12 %)	20 (20 %)	4 (7 %)	3 (10 %)
No	84 (88 %)	80 (80 %)	56 (93 %)	26 (90 %)
Articles in which politicians were named				
Blame	4 (4 %)	0 (0 %)	14 (23 %)	2 (7 %)
Neutral or balanced	14 (15 %)	12 (12 %)	7 (12 %)	10 (35 %)
Credit	6 (6 %)	8 (8 %)	2 (3%)	0 (0 %)
Total articles	95	100	60	29
Article length in words				
Mean [SD]	579 [432]	360 [308]	686 [425]	493 [462]

The German and Finnish results relate to the stratified samples of articles. The British and French results relate to all the relevant articles found in the searched newspapers.

Table 3. Characteristics of the Countries' School Systems in 2007.

	Germany	Finland	Britain	France
Control of education system	Responsibility of Länder. Little direct central control. Regional Education Ministers Conference (KMK) provides some standardisation.	Run by local government with central regulation and influence (resulting in quite uniform system)	England and Wales: Run by local authorities with central control and funding	Centralised, responsibility of Minister of Education.
Compulsory nationwide standardised tests (before age 16)	No (developing)	No	Yes (England) No (Scotland and Wales)	Diplôme National du Brevet (at age 15)
Tests used to rank schools	No	No	Yes (England)	No
National curriculum	No	No	Yes	Yes
Types of schools (to age 16)	Selective after age 10 - 12 (minority at non-selective schools depending on region)	Non-selective	Non-selective (>80% of pupils, remainder in selective state or independent schools)	Non-selective to age 15, selective thereafter.
Salience of education ^a	18 %	4 %	8 %	9 %

^a Percentage of citizens placing education in top two issues in Nov-Dec 2007 (EU average = 9%) (Eurobarometer 68 http://ec.europa.eu/public_opinion/archives/eb/eb68/eb_68_en.pdf)

Table 4. Newspapers included in the 2007-8 search with their 2008 circulation figures and the number of articles found.

Country	number of articles in sample	Newspaper	Daily Circulation^a
Germany	4	Bild, with Bild am Sonntag	3,328,000
	14	Süddeutsche Zeitung	448,000
	20	Frankfurter Allgemeine Zeitung with FAZ am Sonntag	334,000
	21	Die Welt, with Welt am Sonntag	278,000
	17	Taz	56,000
	19	Rheinische Post Düsseldorf	383,000
	Totals (Germany)	95	
Finland	32	Aamulehti	139,000
	12	Helsingin Sanomat	412,000
	9	Iltalehti	123,000
	9	Kaleva	82,000
	5	Keski-Suomalainen	74,000
	33	Turun Sanomat	112,000
Totals (Finland)	100		942,000
France	10	Le Monde	317,000
	9	Le Figaro	328,000
	3	Libération	132,000
	2	Aujourd'hui en France	523,000
	4	Ouest France	794,000
	1	Sud Ouest	319,000
Totals (France)	29		2,413,000

Britain	10	Daily Mail with Mail on Sunday	2,350,000
	6	The Times with Sunday Times	638,000
	5	Daily Telegraph with Sunday Telegraph	885,000
	14	Guardian with Observer	360,000
	5	Herald (Scotland)	78,000
	20	Western Mail (Wales)	38,000
Totals (Britain)	60		4,349,000

Sources: Germany: Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V. (IVW), 2nd quarter 2008; Britain: Audit Bureau of Circulations, June-Dec 2007; France: Association pour le contrôle de la diffusion des médias (OJD) for the national press and Syndicat de la Presse Quotidienne Régionale (PQR) for the regional publications (2007 data) ; Finland : Finnish Audit Bureau of Circulations <http://www.levikintarkastus.fi/levikintarkastus/tilastot/Circulations2008.pdf>. Circulation figures refer to daily editions.

^a The combined circulations of the sampled newspapers were 6, 7, 4 and 18 per cent of the populations of Germany, Britain, France and Finland respectively (2008 figures).

Figures

Figure 1. Number of relevant articles from our search of German, Finnish, British and French newspapers in 2007-08. See Appendix for details of search terms and newspapers.

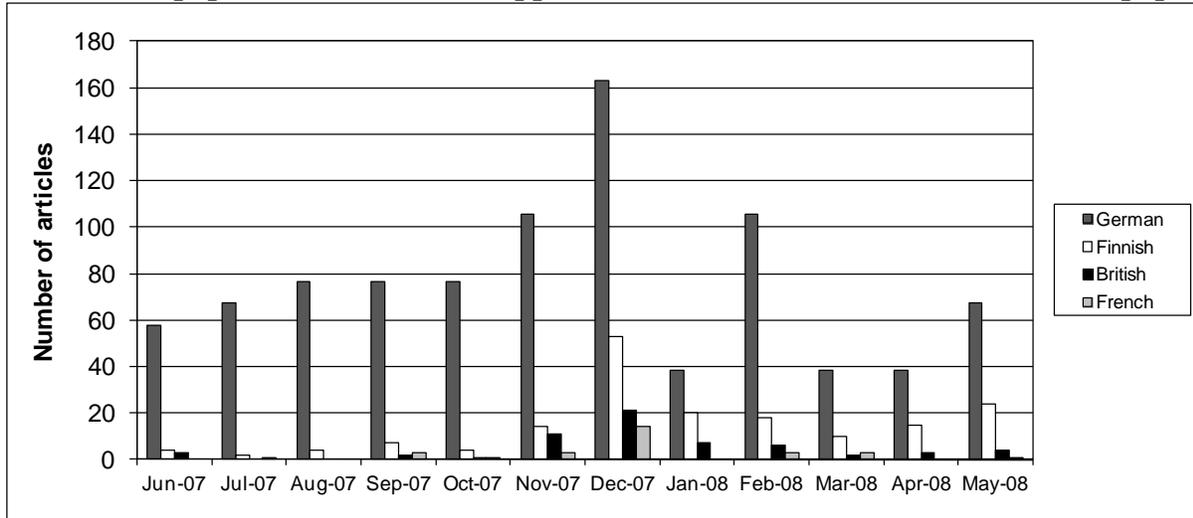


Figure 2. Number of negative, neutral and positive articles in the coded samples by country and date.

