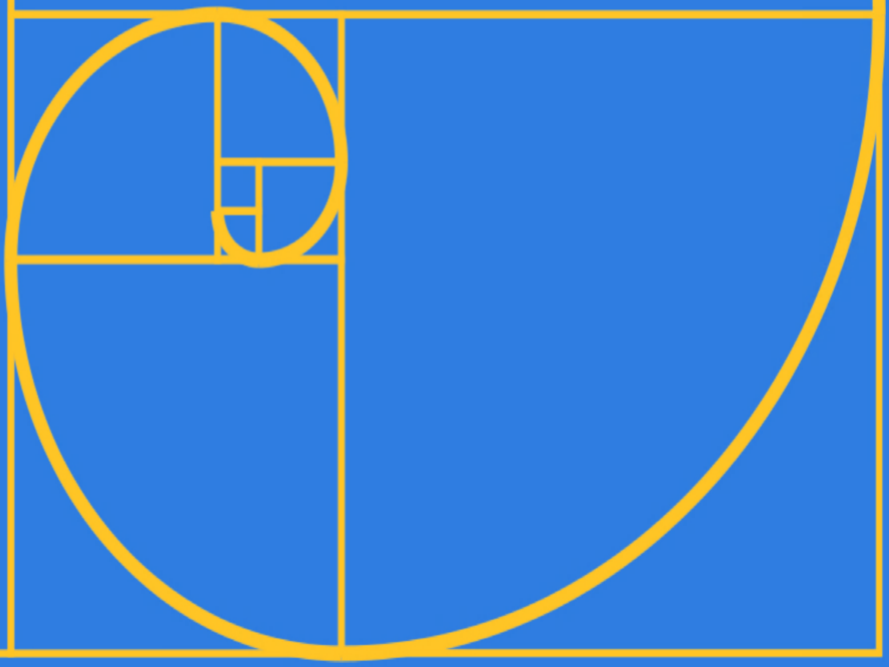
Letnik XXV, številka 5–6, 2014
Revija za teorijo in raziskave vzgoje in izobraževanja

Šolsko polje

Evidence from the PISA Study on Educational Quality in Slovenia and Other Countries

ed. Mojca Štraus



Šolsko polje

Revija za teorijo in raziskave vzgoje in izobraževanja
Letnik XXV, številka 5–6, 2014

Solsko polje je mednarodna revija za teorijo ter raziskave vzgoje in izobraževanja z mednarodnim uredniškim odborom. Objavlja znanstvene in strokovne članke s širšega področja vzgoje in izobraževanja ter edukacijskih raziskav (filozofija vzgoje, sociologija izobraževanja, uporabna epistemologija, razvojna psihologija, pedagogika, andragogika, pedagoška metodologija itd.), pregledne članke z omenjenih področij ter recenzije tako domačih kot tujih monografij s področja vzgoje in izobraževanja. Revija izhaja trikrat letno. Izdaja jo *Slovensko društvo raziskovalcev šolskega polja*. Poglavitni namen revije je prispevati k razvoju edukacijskih ved in interdisciplinarnemu pristopu k teoretičnim in praktičnim vprašanjem vzgoje in izobraževanja. V tem okviru revija posebno pozornost namenja razvijanju slovenske znanstvene in strokovne terminologije ter konceptov na področju vzgoje in izobraževanja ter raziskovalnim paradigmam s področja edukacijskih raziskav v okviru družboslovno-humanističnih ved.

Izdajatelja: Slovensko društvo raziskovalcev šolskega polja in Pedagoški inštitut

Tisk: Grafika 3000 d.o.o., Dob
Naklada: 400 izvodov

Revija *Šolsko polje* je vključena v naslednje indekse in baze podatkov: *Contents Pages in Education; EBSCO; Education Research Abstracts; International Bibliography of the Social Sciences (IBSS); Linguistics and Language Behavior Abstracts (LLBA); Multicultural Education Abstracts; Pais International; ProQuest Social Sciences Journal, Research into Higher Education Abstracts; Social Services Abstracts; Sociological Abstracts; Worldwide Political Science Abstracts*

Šolsko polje izhaja s finančno podporo Pedagoškega inštituta in Javne agencije za raziskovalno dejavnost Republike Slovenije.

Tiskana izdaja: ISSN 1581–6036
Izdaja na zgoščenki: ISSN 1581–6052
Spletna izdaja: ISSN 1581–6044

Letnik XXV, številka 5–6, 2014
Revija za teorijo in raziskave vzgoje in izobraževanja

Šolsko polje

Evidence from the PISA Study on Educational Quality in Slovenia and Other Countries

ed. Mojca Štraus

Contents/Vsebina

1 EDITORIAL/UVODNIK

The Timeless Questions About Educational Quality

Mojca Straus

Concerns about the quality of education systems have now substantially shifted from the questions about the quantity and quality of resources, such as school buildings and accessibility, to the questions about the outputs of the educational process, such as student achievement. Achievement has become one of the key indicators used in evaluating the quality of education systems. Furthermore, these questions are not constrained to the local contexts but are globalized in the sense that the outputs of educational systems, working in different societal and economical contexts, are compared. To address the comparative information needs in the process of the educational quality control, several large scale assessments have been launched in the last decades primarily to provide an information base from which the hypotheses about stability and change in education can be tested.

Some questions remain the same throughout these decades, such as how well do students in a particular country perform in comparison with students from other countries, do they reach expected levels of achievement and what should be expected of them. Further questions pertain to the methodology of the studies and validity of the usage of their results in more general contexts. The quality of an education system proves to be a complex concept that needs constant attention at all levels of the system. This thematic issue is devoted to findings emerging from the latest cycle of the Programme for International Student Assessment (PISA). Slovenian and foreign authors present views and reactions to the PISA methodology and results in the efforts for assessing the quality of the education systems in their respective countries.

In his article, Darko Štrajn discusses relatively recent criticisms of PISA. The criticisms focus on the ranking of results that inscribe PISA as the foun-

dations of the neoliberal market competition entering the education field. Since the initiation of PISA, there were many discussions about the impact of the study's ranking results on the educational policies and processes around the world. In his analysis, the author explores different views on these impacts on the general understanding of the meaning and working of education as well as educational policy development.

In the next article, Urška Štremfel addresses the impact of Slovenia's below average results in reading literacy on the country's educational policy. Using policy analysis, the author provides insight into the first steps of the process of improving the Slovenia's PISA results. The author discusses the importance of having nationally defined educational priorities and goals in order to be able to actually derive a well-defined policy problem and to find the appropriate policy solution to this problem, for example by drawing lessons from the successful results of other participating countries.

One of the countries from which Slovenia might decide to learn from is Germany. In the last twelve years since the first, so-called PISA shock in 2000, Germany has successfully improved its PISA results. In their article, Christine Saelzer and Manfred Prenzel describe three major aspects of Germany's educational development; a thorough diagnosis of the problems in the country's educational system, an intense discourse between all relevant actors, and the implementation of nationwide, overarching programmes to improve teaching and learning. These elements and their impact on German students' PISA results are analyzed. Based on the PISA 2012 results, it is evident there has been a positive educational development in Germany.

Another country, Canada, has been considered very successful in PISA since its beginning in 2000. However, the recent downward trend in the country's results have initiated the call for action. The issues around Canada's PISA results and reflections of different educational actors are presented by Pierre Brochu. The author analyzes the important considerations in the efforts of finding the appropriate levers for changing the observed negative trend in Canada's student achievement.

In the United States of America, the educational policy is developed at the state level. Maria Stephens and Anindita Sen address considerations arising when three U.S. states – Connecticut, Florida, and Massachusetts – derived comparisons of states' results from the PISA data as well from data of other international studies. When different assessments sometimes indicate different or even contradicting results about the educational quality, the important question is what specific factors might explain the observed differences.

Ana Kozina and Ana Mlekuž studied the relationship between PISA 2012 mathematics achievement and attribution styles. In their article, they use national as well as international perspective for investigating students' attributions of causes for success and failure on the PISA 2012 mathematics achievement test in relation to actual test score. They conclude that attribution for success should be considered in educational setting for example in communicating praises for students' success in a manner promoting effort.

In the final article, Mojca Štraus explores the roles of socio-economic background and mathematics-related attitudinal factors in explaining achievement in mathematics literacy of the PISA 2012 study for Slovenia in comparison with Germany, Canada and the United States. Mathematics-related self-beliefs are shown to be stronger predictors of achievement than students' drive and motivation and similarities are observed between the Slovene and German students' responses as well as between the Canadian and the United States students' responses.

The articles in this issue show that data from international assessments of student achievement represent a rich source of information on education systems in the world. However, thorough understanding of the design, methodology and implementation of the assessments is of vital importance for making valid and useful interpretations of the results. The general steps in conducting an international comparative assessment are that participating countries agree on the population of students and the curriculum domain to be assessed, and on an instrument to assess achievement in the chosen domain. The instrument is administered to a representative sample of students in each country and comparative analyses of the data are carried out. These analyses are intended to provide information about the educational quality in the form of comparisons of students' scores or sub-scores on an international achievement test. An important part of this is understanding the reasons for observed differences between and within the countries from the collection of the background data, especially in the areas where weaknesses in achievement are identified.

The story of educational quality control does not end with the publication of the international or national comparisons of participating countries' results. After the PISA 2012 results were published in late 2013, countries started with additional qualitative and/or quantitative studies designed to unravel the origins of the observed weaknesses in order to set up and carry out the appropriate remedial actions. The contributions in this issue show examples of such analyses and the interpretations of the findings. It is shown that the internationally comparative data are most often used for the functions of descriptive comparisons and trend analyses

and that it is more difficult to provide answers about the causes of any observations. When the results of the studies are used, one needs to be careful in drawing conclusions. There is an abundance of caveats that could diminish the validity of these conclusions ranging from the sampling design and response rates, coverage of domain and instrument design, data collection procedures, motivation of students and other respondents, technical procedures in data analysis and, not least, inappropriate causal inferences.

As evident from the contributions in this issue, the data from an assessment of student achievement do not, by themselves, convey messages about the quality of education or evaluation of the reforms that have been implemented. The data collected need to be interpreted with a reference to relevant comparisons, for example to the goals of education in a particular country or to the results of other countries. However, setting absolute standards in education is difficult. To try to set realistic standards for educational system comparisons with other relevant countries are essential. As shown in this issue, this is important in the largest education systems in the world and even more so in Slovenia.

The overall problem with analyses of the assessment data is how to address the imminent questions on the educational quality and effectiveness without reporting information that is easily misunderstood and/or misused. It is very difficult to determine abundant factors within or outside the education system that influence achievement. Moreover, conclusions from an assessment rarely offer clues about causal inferences. They can, however, be useful as circumstantial support for the conclusions about the determinants of achievement or as a source of inspirations for finding possible levers of improvement in further research. There are, nonetheless, important reasons for the usefulness of such studies. Not unimportantly, assessments are relatively inexpensive compared to other aspects of managing education, such as implementing curriculum changes that involve substantial professional development of teachers. Further, it is easier to mandate assessment requirements at the system level than it is to take actions that involve actual change in what happens inside the classroom. Such studies are therefore useful for getting the overall picture of the status of the things in education. And, as a consequence of media attention given to the international assessments, international studies can help education to become a priority among the areas that need policy makers' attention.

2 PAPERS/RAZPRAVE

The PISA Syndrome: Can we Imagine Education without Comparative Testing?

Darko Štrajn

Introduction

Unlike "normally" by citing academic books and journals – I am starting this article by recalling relatively recent criticisms of OECD's PISA testing addressed to wider public. The fact that this criticism is recent does not imply that it is also entirely new. The logic of this criticism, which has been detectable almost ever since the inception of PISA – and indeed since much earlier pioneering IEA studies like FIMS, TIMSS, and so on in more than just governmental settings – had been conducted, has gone public on a grand scale. *The Guardian*, Tuesday 6th May 2014, published a letter addressed to PISA director Dr Schleicher under the title "OECD and Pisa tests are damaging education worldwide." The letter was signed by many distinguished academics from universities (mostly American and European) and some other interested public personas. This academic public gesture had a quite strong echo in world press. However, answers by the PISA director and by members of a global network, consisting of researchers, who actually work on designing and implementing PISA testing, were much less published in the world press. Another case of recent public criticism of PISA is Erwin Wagenhofer's film documentary *Alphabet* (2013), which actually commences with a strong point on how educational achievements of Shanghai schools were under the influence of PISA testing. The type of education, which is adapted to achieving high scores in PISA testing, especially in the fields of mathematics and natural sciences, presumably – as it is stated at the beginning of the film – flattens children's creativity, ability to think critically and independently. Both of these critical statements aimed at policy makers, and even more to the broader public, expose what they see as a dubious nature of

ranking of results that inscribe PISA into the foundations of the neoliberal extension of market competition to all avenues of life. However, exactly the rankings, as they are presented in league tables in a somewhat quick succession once in three years, made PISA so "popular" and influential. Therefore, any abandoning of such presentations of the results seems quite unimaginable. On the other hand, a dilemma on whether these rankings are consequences or causes of what has been seen as educational transformation in favour of global neoliberalism seems pertinent, but hard to answer.

In this paper, I shall just briefly discuss the main lines of argument in the above mentioned public outcries against PISA and in the next step I shall take a look at some examples of academic deliberations on PISA testing. Further on, I will be exploring on the paradigmatic level for "deeper" reasons for such disputes and insuperable differences, concerning cultural, methodological and theoretical aspects of these considerations. At the end of the paper, I shall try to open questions on how PISA testing nevertheless makes sense.

Questions and Answers

The views, which are expressed in *The Guardian* letter (Andrews, 2014), represent an important step in discussions about standardised testing precisely because they are communicated to a larger public. This means that we can take them to be an attempt to make an impact on public policies, as well as trying to influence a critical understanding of such procedure as PISA testing and its results. In all fairness to the signatories' good intentions, it should be noted that they do not *a priori* reject the very method of testing itself and, in spite of the rather harsh criticism; they give suggestions on how PISA should proceed in its work to attain socially and educationally more acceptable impact. The signatories assert that PISA "/.../has contributed to an escalation in such testing and a dramatically increased reliance on quantitative measures," which has, in their view, resulted in many negative effects. Just three years assessment cycle shifts attention to short-term policies, which are mostly inappropriate in various cultural contexts. PISA is further, in the signatories' opinion, too focused on measurable aspects and so it "takes away attention from the less measurable or immeasurable educational objectives." PISA is then, among other problematic effects, blamed for an increase of "public-private partnerships," which sustain for-profit educational services in America and project them also in Africa. After avowing some more harmful consequences of PISA, such as it is conducted for last 13 years, the authors of the *Guardian* letter make seven "constructive ideas and suggestions." Since my intention is

not to deal with the whole spectrum of problems, which these "ideas and suggestions" touch upon, let me only mention that the first suggestion requires from OECD to "develop alternatives to league tables" and to "explore more meaningful and less easily sensationalised ways of reporting assessments outcomes." The letter is concluded by questioning the legitimacy of OECD as an organisation for becoming a "global arbiter of the means and ends of education." The authors of the letter find that the "international competition for higher test scores" harms diversity among cultures and traditions.

A direct answer to these allegations under the title "OECD's PISA under Attack!" signed by almost 400 above all "researchers of school performance" (as they chose to present themselves) from all continents is without any doubt an illustration of the fact that the academic sphere is divided on most questions raised in *The Guardian* letter. Of course, I have no intention to judge who is right in this dispute. The answer to *The Guardian* letter is obviously an upshot of a quite quick reaction. Therefore, the answer mainly succeeds in demonstrating that, at least, there is a strong misunderstanding on the matter between members of research communities, which are supposed to know what is there to know about the testing of school achievement. Still, I would dare to say that the answer seems somewhat weak. It essentially boils down to this assertion: "PISA student assessments, like other similar kinds of tests around the world, have the same function of a thermometer in medical diagnostic." (Ichino, 2014) We can take this as a statement on PISA being essentially just a "neutral" instrument. The medical metaphor, which is further elaborated, seems to be unsatisfactory as an answer.[1] Beside this, as it appears to me, the answer imputes to *The Guardian* letter an intention, which it did not have, saying that it was "clearly aimed at excluding comparable evidence of student performance from educational decision-making." The "coming out" into the open public space of the two academic groupings points towards a need to rethink the role of PISA testing not only in order to fight social battles in the academic arena, but also in order to distinguish between research results and its (ab)uses, and then to at least recognize differences in justifiable approaches to such complexities as educational in-

1 The signatories of the answer to *The Guardian* letter probably meant to address not just the academic community and, therefore, they picked a linguistic short-cut to readers. Still, it should be pointed out that metaphors can be tricky. Let me cite just one example of many similar notices (of which early examples can be found also in Plato's Socrates dialogues): "Metaphor is helpful (and even indispensable) as vehicle to think about abstract phenomena, but one should be careful not to mistake the metaphors for the 'reality' they try to describe." (Boers, Demecheleer, 1997, p. 116). It is also interesting that medical metaphors are very much used in many discourses on economy.

stitutions. This would make possible to tell apart intellectual, social and political phenomena from genuine research problems.

Another example of recent criticism of PISA, aimed at larger public in the form of the movie *Alphabet* (2013), can be taken as an interesting case of opening the eternal question of goals and senses of education. I tend to agree with those observations of the film, which see its grasp of education in today's global world as a bit simplistic, pretentious, biased and even misleading, but still the movie could be commended in its main intention to sound an alarm about current developments in education and its role in globalisation. Still louder sounds the alarm, which the film raises in view of the forms of domination on the level of social practice in corporate management.

> "Wagenhofer's actual beef appears to be not with schools but with the system itself, which emphasizes bloodthirsty, profit-driven competition over the prenatal connection humans feel to their mothers. With apparent alarm, the film cites studies showing that people lose their capacity for 'divergent thinking' over time, which, it doesn't take a divergent-thinking genius to realize, necessarily follows from standardized education." (Debruge, 2014)

This perceptive observation, taken from the film review, published in one of the most prominent film magazine, applies to the problems and paradoxes, which PISA could not avoid even if it tried no matter how hard. As a part of the activities of OECD in the field of education, the whole structure of PISA is having a stable support and necessary institutional authority, but this also brings about suspicions of apparent adjusting of the research profile to the broader politics of this intergovernmental organisation such as OECD is. Declarations by PISA advocates that the testing, as it were, happens to be "just a neutral instrument" rouses ceaseless arguments about the ethics of research, which concerns social research even more than the research in natural sciences, since the effects of the results might be hypothetically more complex and prone to manipulation. The rankings apparently generate various kinds of competitions within and between countries and in a "trickle down" effect strengthen debateable "neoliberal" socialisation of youngsters. However, at the same time PISA produces a huge amount of varying data, which many researchers, independently of their political views, find almost indispensable. Unfortunately, politicians and policy makers see their usages in their own way, which the researchers cannot always control.

PISA, Neoliberalism

Anyway, many of these aspects were and keep being discussed in the global research community in less publicly exposed, but nevertheless strongly controversial discourses. Many disputes, divergent studies, books and articles predominantly in less agitated discourses ponder the social role, impacts, advantages and shortcomings of PISA and also of other similar assessments of education, done by methods of testing; many doubts are raised as well about the benefits of rankings and benchmarking, as consequences of testing. Other aspects of debate touch upon the impact, which PISA has on the structure of the curriculum, for instance, in a direction of stronger emphasis on one type of knowledge at the expense of the other: favouring natural sciences and mathematics and diminishing the importance of humanities and critical thinking. Publications concerning PISA are, of course, abundant, but one can quickly discern between those studies, which more or less take the results of PISA tests for granted and use them in order to come to terms with what is going on in educational systems and those discourses, which take a critical distance and observe in various degrees of criticism ostensibly worrying effects of PISA. These criticisms cannot be easily typified, but they are mainly based on similar, albeit much more elaborated, theses as the main points of *The Guardian* letter. With a dose of simplification one can say that a part of world's researchers in the field of education and a number of scholars, mostly from humanities, take PISA to be above all an agency of globalisation along the lines of global capitalism and its neoliberal ideology. Many critical authors would agree with such propositions as this: "When we speak of neoliberal *policies* throughout the world, it is not only because they exist in the platonic world of ideas or only because they constitute a space of possible options, but also, and perhaps above all, because we put some of them *into action*, and they are followed by effects." (Hilgers, 2013, p. 78) Further on, similarly to Joel Spring, many authors are increasingly naming the bearers of these options: "Neoliberalism is an important part of educational discourses in IGOs, such as the World Bank, OECD, and WTO, and within national governments." (Spring, 2008, p. 343) Propagation of (curriculum and/or culture) uniformity and unfair competition, which is induced through rankings of countries according to a level of students' "success", seems to be the most frequent reproaches. "At the school level, Program for International Student Assessment (PISA) is the best known example of international rankings and is an interesting example of how a transnational organization such as the Organization for Economic Cooperation and Development gains influence in different ways over national educa-

tion reforms in both its member and non-member countries." (Parmenter, 2014, p. 203) I cannot present here the whole spectrum of such criticisms, of which some happen to be quite sophisticated and many of them would probably deserve at least the benefit of a doubt also from PISA designers themselves. As I mentioned in the introduction to this paper, critical attitudes in comments on PISA are not new, therefore the signatories of *The Guardian* letter could easily point to a large basis of theoretical arguments on which their argument was built. One of the most heard voices among the critics of the tendencies in education policies in the German-speaking world belongs to the Austrian philosopher of education Konrad Liessmann, who became especially upset because of the PISA rankings in league tables. "All relevant and also publicly widely debated decisions of educational policies from last years are either motivated by an inferior position on the league table or by a wish to attain a better position on the list." (Liessmann, 2006, p.74)[2] Liessmann's points, of course, do not end with this. His whole argument concerns the confronting of all them agents of the neoliberal world and education, such as it has been conducted after recent changes of curriculum and school management styles, to the tradition of the Enlightenment and goals of education, as they are comprised in the notion of *Bildung*, which is characteristically almost untranslatable to English.

> "Instead of the educational aims of the Enlightenment – autonomy, self-consciousness, and spiritual comprehension of the world –, instead of the educational goals of the reformist pedagogies – real-life orientation, social competence and joy of learning –, instead of the educational goals of the politicians of neoliberal school – flexibility, mobility and employability – there is only one educational target: to withstand PISA!" (Liessmann, 2006, p. 75)

It should be noted that Liessmann's observation ascribes to PISA that it even transcends neoliberalism and its social aims by narrowing its focus just on competition.

Written not much earlier, the book of Christian Laval had a large echo and public impact in the French-speaking world, the book claiming that "school is not an enterprise", which analysed the "neoliberal attack on public school." It goes without saying that in Laval's criticism, PISA is blamed for its contribution to the cult of efficiency, for the practice of benchmarking and for culture of evaluation as a system of control. (Laval, 2004) Mojca Štraus and Neja Markelj represent a different case of indicat-

2 Since Liessmann's work isn't translated in English both citations in this article are my own translations from German. Therefore, I am accountable for anything that gets lost in translation. The same goes for other translations of citations from Slovene and French in this paper.

ing the same general change in perceptions of a social role and the meaning of education due to PISA. In the context of their study on what PISA results could mean for the decision makers in Slovenia, they wrote: "Orientation to the development and measuring of competencies seem to be a reflection of the emphasis on the function of education as the production of human capital. Relocation of emphasis from knowledge to competences can also be seen as an example of the aforementioned research to support decision-making in education." (Štraus, Markelj, 2011, p. 37) To conclude this part of examining examples of criticism of PISA testing, let me cite a bit longer fragment, which confirms the point on the difference in perceptions of PISA.

> "PISA results are frequently discussed and debated in the policy world and among education researchers. While PISA supporters paint a bright picture of PISA and how it can bolster education in today's globalized world, its critics draw attention to the negative consequences of PISA. Education has, thanks to PISA, moved away from the enlightenment ideal of promoting personal development and creating reflective and culturally aware citizens, towards an ideal of education in the interest of economic growth, promoting performativity, standardization, and decontextualization – according to some of its critics (cf. Carvalho, 2012; Lawn, 2011; Mangez and Hilgers, 2012). Advocates of PISA do not consider this shift negative. On the contrary, benchmarking education systems and testing the life skills needed in today's world are claimed to be a great help, informing policy for education system development (Schleicher, 2013)." (Hanberger, 2014, p. 2)

These observations bring us to a question of causes and effects. Did PISA cause the advancement of neoliberal politics into the sphere of education or did the complex development of neoliberal capitalism open the research space for PISA? Is a shift from evaluating knowledge to testing competencies "restructuring" school as an institution and its complex role in any society? However, while discussing criticism of PISA, one cannot avoid worries, expressed in a different register. From the "epic" times of the first few cross-national studies, which were conducted by *The International Association for the Evaluation of Educational Achievement* (IEA), a threat of cultural homogenization was indicated by many writers. PISA, which stepped quite a bit later into the amphitheatre of international assessment of school achievement, only strengthened such fears.

PISA, Culture

What has just been said unties a little bit the strictness in the relationship between PISA and the notion of neoliberal capitalism, since the fears

around the cultural impact of cross-national assessments of education appeared already at the time, when neoliberalism was just an obscure theory, cultivated by a group of scholars, economists and some philosophers, who joined their ranks under the name of Mont Pelerin Society. Of course, many reflections on PISA are enunciated in the context of post-colonial studies, gender studies and other contemporary forms of critical thinking that are often associated with political anti-globalisation movements, which also include a range of alternative education practices and experiments. However, I am not entering in these interesting debates since their stress on complexities and sometimes their attention to details, exceed the main focus of this paper. Although the kind of criticism that brings up problems of cultural impact is much more multifarious than just the criticism of "PISA's neoliberalism," there is a starting point, which could be expressed, as follows:

> "What are the politics and sociology and anthropology of the international testing movement as if 'educational results' were a sporting event? The second comparative puzzle, which attaches to PISA is: in what sense is it 'comparative education'? At what point do numbers become or represent or stand for cultures, and what needs to be explained about the cultures/numbers symbiosis?" (Pereyra et al., 2011, p. 3)

However advancing from such points, opinions get increasingly different. Obviously, more than establishing any firm evidence of PISA's transforming impact on cultures, PISA represents a reference point, which arranges quite a number of discourses on a relationship of culture and education in our complex world. South Korea was always excelling in (not only in PISA) cross-national schools assessments and at the same time educators there seem to be "culture sensitive." Surprisingly, the Korean critic sees as a threat exactly that educational tendency, which in view of most PISA critics is more supressed than promoted by testing.

> "In this tendency toward individualized and differentiated educational processes that are assumed to foster students' creativity and independent thinking, it is natural to criticize 'traditional' Korean education, which is portrayed to have limited students' exposure to individualized and differentiated curricula and instruction. However, as I have shown in my work, the recent educational reform for individualized and differentiated education has actually reduced the strength of 'traditional' Korean education, which helped low achieving and socioeconomically disadvantaged students maintain a comparatively high level of academic achievement compared to corresponding students in other countries." (Ilyunjoon, 2014, p. 3)

Perceptions of PISA's "cultural impact" actually vary since most authors are aware that there are other agencies of a global "cultural homogenization" that might have benefited from PISA, which indeed tends to be "culture-blind." Educational systems and their elements (like curricula, teaching methods, school management, and so on) of course change, and, of course, they are always making part of cultural context. "/.../for many countries in the world that has happened is a shift in what could be called the topography of education. Between the early nineteenth century and the early twenty-first century, the map of 'education' itself has changed. Its contents, its institutions, and the people who populate it have been reconfigured." (Cowen, 2011, p. 30) A quick "meta-analysis" of PISA impacts would probably show that educational systems still conform to their local social and cultural contexts – which are in their turn changing either in a progressive or conservative direction – in spite of responding to some "incitements" from PISA results. China's case is typical in this respect.

> "/.../ our analysis of the reasoning surrounding the PISA results reveals that there is a profound discrepancy between local political actors and stakeholders on the one hand and independent researchers and overseas professors on the other. The discourse centring on the PISA 2009 results has reshaped the education discourse in China. The case of China is particularly interesting for education discourse analysis, because the pre-PISA discourse had been characterized by the criticism of the exam oriented education and the scepticism of the effectiveness of the education reform." (Zhang, Akbik, 2012, p. 26)

I am leaving many other aspects of the "cultural problem" of PISA open, since the above-mentioned facets are maybe sufficient to exemplify the type of the problem.

Paradigmatic Divide

Epistemological questions will always represent issues for differences among researchers. Such questions, of course, open problems of methods, which are unavoidably intellectually funded. Undoubtedly "the syndrome" of PISA consists of many components. As we can gather from many debates, these components are: conceptual differences, political perceptions, and cultural contexts. However, fundamentally PISA is linked to knowledge as is any education-related phenomenon, which means that it cannot avoid paradoxes of "knowledge about knowledge." Philosophy for centuries searched for a universal model of knowledge. Hence, at least two broad different "paradigms" of reflexive knowledge persist. Philosophers – of course with immense number of nuances – basically agree that these different paradigms could be identified as a difference between em-

piricism and rationalism from 17th Century, or as the difference between positivism and transcendentalism (or constructivism), or as the gap between Anglo-American philosophy and Continental philosophy. Some would also argue that the split between the two basic paradigms is rooted in Antiquity – for instance in the unfinished dialogue from Plato, *Parmenides*, which left readers with unanswered questions on the relation between the part and the whole – others, would see this split in mediaeval logics, and so on. In modernity and postmodernity, there were many attempts to overcome the divide, but it looks that such attempts mostly contribute to just new elaborations of the rift. One of the modern manifestations of the divide – between positivism and deconstruction – was highlighted by Stanley Cavell, who certainly made a few steps towards creating a field of mutual understanding.

> "And I cite their [positivism's and deconstruction's] claims to what may be seen as the discovery of the originariness of writing over voice, of system over individual intervention, of sign over word – since the appeal to mathematical logic for its algorithmic value is an appeal to its sublime inscriptional powers (of alignment, rewriting, iteration, substitution, and so on). Positivism's inscriptionality may be seen as in service of a homogenization of the field of sense." (Cavell, 1994, p. 83)

Cavell's success in bridging the gap between two "universes" of thought made a strong impression in such fields as culture, or, to be more precise, in film theory, as well as in some trends in philosophy itself. We are still waiting for "a Cavell" in the realm of the scientific mind. As it is well known, "positivism" is closely associated with (positive or "exact") sciences. Especially thanks to recent possibilities to acquire and manage large amounts of data, positivism is also re-occupying the space of social sciences, which through the work of Durkheim, critical philosophers, existentialists, and so on, was for a long period a domain of thinking about the world in terms of the notion of totality. PISA is just one of the phenomena in research that makes use of the "positivist" methodologies, which carve out their problem field from the social and cultural complexity. Such methodologies, no matter how well elaborated or specific in their founding they may be, lay claim that the knowledge, which they acquire by applying their rules and "tools", is certain as it is firmly "evidence-based." Usually users of such methodologies – viewed as "partial" by a range of anti-positivist critics – do not hesitate to give the "we don't know" kind of answers for any problems, which are considered to be outside of their methodological framework. However, this insisting on a particular insight, "based on facts," is seen as a synecdoche: the way PISA test results are presented strongly suggests that mathematics and natural

sciences stand for entire knowledge, as well as, that such knowledge is crucial for economic development. Of course, such a supposition can probably not be proved, since such categories of knowledge as historic memory and artistic sense have their role in any social system, and they operate within the economy in a broader sense of the word. On the level of theory, the differences will probably never be settled, since anti-positivists will always insist on an attribute of "instrumentality" of such methods as the ones, used in PISA.

This brief and very superficial explanation of the paradigmatic gap can be taken as just one aspect of many reasons for "misunderstandings" between advocates and adversaries of PISA. However, by taking into account such sophisticated aspects of the differences, one can still find data – no matter how much they are seen to be ideologically constituted, or no matter how they represent only a reduced picture of the "reality," and so on – as representing something. Of course, one is free to decide what they represent. Any decisions of actions in changing the profile of a national education depend on complex local contexts. In spite of credible reproaches, regarding what is voiced as "homogenisation," there is always a space in local policies to advocate "good traditions" against mismatched changes.

Conclusion

It is a truism to say that theoretical and practical constituents of education have always been ingredients of larger social movements. They mark conflicting issues in the politically determined power relations in the educational field. In countless discourses, education keeps recurring as a crucial agency of the social emancipation, both from class or gender oppression and from other forms of cultural exclusion, but also as a precondition for self-accomplishment of an individual. A huge intellectual input into developments, processes and events in educational systems is an inherent force of social-educational movements. As an end of neoliberalism is anxiously hoped for, there is a huge helping backlash of emancipatory educational discourses. However, in light of the question on whether PISA is the cause or effect of structural institutional shifts, adaptations in the economy, and so on, another question on the full pertinence of PISA as a main object of such criticism is relevant. Scholarly volumes of books – let alone journal articles and other not strictly just academic publications – that deal with the role of education in social reproduction and in movements for social change are growing almost exponentially.[3] The out-

3 For instance, I myself wrote some fifteen book reviews for *International Review of Education* (Hamburg, Springer) in the past four or five years, which presented studies on relationship

cry against PISA in The Guardian letter is therefore a kind of cumulative effect of the growing bid for emancipatory education, which again strives to return to a composition of educational ideals instead of the aims comprised in more or less utilitarian and technocratic concepts of increasingly visible failure of such neoliberal projections as knowledge society, human capital, and so on.

> "If school has any sense nowadays, it should awaken in all its forms the reason finding out the emancipatory character of knowledge. /.../ Yes, the historical modernity was wrecked within modernism, in which techno-scientific rationality demolished the subject (*sujet*). Let us find anew its initial project in a dialogue between reason and the subject that originated with the Renaissance and the Reformation still alive at Descartes." (Fabre, 2011, p. 42)

Does all this mean that such comparative testing as PISA, as its most outstanding case, becomes obsolete? In spite of all criticism, the answer should be definitely: "No!" It is visible already in *The Guardian* letter that the authors oppose many features of presentations of the results (rankings) and a number of other impacts of PISA, but testing as a relevant research method is not really attacked. In a final analysis the point of the letter boils down – quite like the point of the film *Alphabet* – to an outspoken condemnation of the neoliberal society. OECD is undoubtedly an organisation of governments, which are entangled by the structures and networks of global capital and such "instruments" as PISA are "taking the pulse" (to use the medical metaphor from the answer to the letter, we talk about here) of education, which operates under such a system. Still, there is no reason to doubt that in the framework of complex methodology, PISA does not deliver very interesting piles of different data. For example, in the volume of "overcoming social background" (OECD, 2010) it seems that the PISA team is trying to react to some criticisms from the agents of "emancipatory currents" since it gives very detailed data in the domain, which is crucial for any thinking about a redemptive role of education. Explicit and well presented – even ranking in this case does not seem harmful – data on correlation between socio-economic background and the performance of students and schools, make it possible for far reaching conclusions. The same goes also for a number of other such reports, which follow after the main and controversial presentation of the results of testing. Slavko Gaber and his co-authors start from the example of France,

---

of democracy and education, on problems triggered by the economic crisis, and so on. Ideology of neoliberalism and such consequences as social inequalities are analysed and criticised in these books.

where only 38% of a generation who reach the educational credential, accumulate an adequate cultural capital.

> "/.../today researchers reinforce their claims about the inadequacy of achievements at the national level with the results of the already well-established international comparisons of knowledge. They don't remain only within data, which show, that in New Zealand and Sweden, there are 80% of those, who 'may hope for a good job,' in Finland, 73% in Poland and Hungary about 70%, but they also take into account the research results of PISA and TIMSS, which allow valid performance comparisons of educational systems and empirically lit analysis of national systems." (Gaber et al., 2009, p. 84-85)

Such comments by researchers of education are not very rare. PISA, therefore, makes possible critical analysis, which even runs against its assumed "neoliberal and homogenising objectives." No matter how well any such criticism is founded, no matter how strong its arguments are, it should be recognised that even so the testing and the acquired data make such criticism and its conceptual achievements possible. Of course, one would like to see more dialogue between different "schools" of comparative research, as well as some pondering on the effects of such presentations as, for instance, the league tables, within PISA organisation itself. On the other hand, one should be aware that controversies in as much as possible unrestrained democratic public space generate breakthrough new ideas and social movements. And this holds true whether controversies are resolved or not.

References

Alphabet (2013) Documentary film. Directed by: Erwin Wagenhofer. Germany: Pandora film.

Andrews, P. et al. (2014) OECD and Pisa tests are damaging education worldwide. *The Guardian* [Online] Tuesday 6 May 2014. Available from: http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics (Accessed: 4th August 2014).

Boers, F., Demecheleer, M. (1997) A Few Metaphorical Models in (Western) Economic Discourse. In: Liebert, W.-A., Redeker, G. & Waugh, L. R. (eds.) *Discourse and Perspective in Cognitive Linguistics.* Amsterdam: John Benjamin Publishing Co.

Carvalho, L. (2012) The Fabrications and Travels of a Knowledge-Policy Instrument. *European Educational Research Journal.* 11 (2), pp. 172–188.

Cavell, S. (1994) *A pitch of philosophy: autobiographical exercises.* Cambridge, Massachusetts, London: Harvard University Press.

Cowen, R. (2011) Edging closer to the hero, the barbarian, and the stranger. A note on the condition of comparative education. In: Tröhler, D., and Barbu, R. (eds.) *Education Systems in Historical, Cultural, and Sociological Perspectives.* Rotterdam: Sense Publishers.

Debruge, P. (2013) *Film Review: 'Alphabet'.* Variety [Online] December 3, 2013. Available from: http://variety.com/2013/film/reviews/film-review-alphabet-1200877233/ (Accessed: 5th August, 2014).

Fabre, M. (2011) Éduquer pour un monde problématique. La carte et la boussole. Paris: Presses universitaires de France.

Gaber, S., Marjanovič Umek, L., Podlesek, A., Sočan G. and Tašner, V. (2009) Analiza razlik v dosežkih učencev/dijakov ter analiza primarnih in sekundarnih učinkov družbenih razlik na dosežke učencev/dijakov [Analysis of Differences in Student Performance and of Primary and Secondary Effects of Social Differences]. ***Šolsko polje.*** XX (1-2), pp. 83-125.

Hanberger, A. (2014) What PISA intends to and can possibly achieve: A critical programme theory analysis. *European Educational Research Journal* [Online]. 13 (2), pp. 167-180. Available from: http://dx.doi.org/10.2304/eerj.2014.13.2.167 . (Accessed: 8th August, 2014).

Hillgers, M. (2013) Embodying neoliberalism: thoughts and responses to critics. *Social Anthropology/Anthropologie Sociale.* 21 (1), pp. 75-89.

Hyunjoon, P. (2014) No matter how high your test score is, you are still bad: Korean education's responses to PISA. *Revue internationale d'éducation de Sèvres* [Online], Colloque: L'éducation en Asie en 2014: Quels enjeux mondiaux. Available from: http://ries.revues.org/3749 (Accessed: 11th August, 2014).

Ichino, A. et al. (2014) OECD's PISA under attack! [Online] Available from: http://www.eale.nl/about%20eale/OECD'S%20PISA%20under%20attack!.htm (Accessed: 4th August 2014).

Laval, Ch. (2004) *L'école n'est pas une entreprise. Le néo-libéralisme à l'assaut de l'enseignement public.* Paris: Editions La Decouverte.

Lawn, M. (2011) Standardizing the European education policy space. *European Educational Research Journal.* 10 (2), pp. 259-272.

Liessmann, K. P. (2006) *Theorie der Unbildung. Die Irrtümer der Wissensgesellschaft.* Wien: Paul Zsolnay Verlag.

Mangez, E., and Hilgers, M. (2012) The Field of Knowledge and the Policy Field in Education: PISA and the production of knowledge for policy, *European Educational Research Journal.* 11 (2), pp. 189-205.

OECD (2010) *PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Volume II).* Paris: OECD publishing.

Parmenter, L. (2014) Globalization in Japan: Education Policy and Curriculum. In: Nelly, P., Stromquist, N. P. & Monkman, K. (eds.) *Globalization and Education / Integration and Contestation across Cultures - 2nd Edition.* Lanham, Boulder: Rowman & Littlefield.

Pereyra, M. A., Kotthoff, H.-G. & Cowen, R. (2011) PISA under examination: Changing knowledge, changing tests, and changing schools: An Introduction to the Comparative Puzzle. In: Pereyra, M. A., Kotthoff, H.-G. & Cowen, R. (eds.) *PISA under Examination.* Rotterdam: Sense Publishers.

Spring, J. (2008) Research on Globalization and Education. *Review of Educational Research.* 78 (2), pp. 330-363.

Schleicher, A. (2013) *'Use data to build better schools'.* TED-conference published 21st February 2013. Avaliable from: https://www.youtube.com/watch?v=7Xmr87nsl74 (Accessed: 26th July 2014).

Štraus, M., and Markelj, N. (2011) Bralna, matematična in naravoslovna pismenost dijakinj in dijakov 1. letnikov srednjih šol v Sloveniji v raziskavi PISA 2009 [Reading, mathematics and science literacy in 1st class of upper secondary school students in Slovenia in PISA 2009]. Šolsko *polje.* XXII (5-6), pp. 35-68.

Zhang, Ch. and Akbik, A. (2012) PISA as a legitimacy tool during China's education reform: Case study of Shanghai. *TranState Working Papers*, No. 166 [Online]. Available from: http://hdl.handle.net/10419/64810 (Accessed: 4th August 2014).

Slovenia on its Own Way Towards Improving PISA Results

Urška Štremfel

Introduction

PISA (Programme for International Student Assessment) becomes a prevalent assessment of the national education systems in the last decade (Hopmann et al., 2007; Pereryra et al., 2011; Meyer and Benavot, 2013). PISA results, presented in comparative achievement scales, provide an insight into how one educational system performs in comparison to other systems and also how one educational system contributes to the achievement of common goals of particular group of participating countries (e.g. European Union (EU) member states together decided a benchmark to have less than 15% of low achievers[1] in PISA by 2020) (Council of the EU, 2009). Since PISA results and results of other international comparative assessment studies[2] often becomes incorporated in the national educational targets, PISA also helps to identify how successfully participating countries follow their national priorities and goals.[3] There is one additional insight that PISA allows. The design of PISA, which is conducted in cycles, enables the monitoring of changes in students' outcomes over time. Such changes indicate how successful education systems have been in developing the knowledge and skills of

1 PISA provides a profile of students' performance using six proficiency levels. The low-achievers are students, who do not reach the proficiency level 2, which present a baseline level of literacy at which students begin to demonstrate the competencies that will enable them to actively participate in life situations (OECD, 2010a).

2 E.g. Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS).

3 E.g. Slovenian White Paper on Education (2011, p.25) states "At the state level we need to state and map out a clear path towards the goal, that performance of Slovenian students in international comparative assessment studies are at the top, that mean at least in the upper third of the students' achievement of the developed countries".

15-year-olds. All countries seeking to improve their results can therefore draw and learn lessons from those that have succeeded in doing so since 2000, when the PISA was first conducted (OECD, 2010a, p.13).

The importance that PISA has gained in the assessment and development of national educational systems is often understood in terms of transnational policy making (Meyer and Benavot, 2013). If we understand the policy making as the solving the policy problems of / for society (Lasswell, 1951), we can also argue that it can be understood as transnational problem solving (Scharpf, 1997). That means that PISA helps participating countries to understand the weakness of their national educational systems (in international comparative perspective) and also provide the environment for finding the right solution of perceived problem. Despite some theoretical reservations towards considering comparative achievement scales as the legitimate source of policy making (e.g. Kodelja, 2005) and exploiting their results for politically motivated changes at the national level (e.g. Štremfel, 2013), PISA has become widely accepted that these comparative achievement scales (called also league tables) present an important source of the identification of national policy problems and finding policy solutions in participating states (see e.g. Grek, 2010). As such comparative achievement scales, if appropriately used, can present an important source not only for the assessment, but also for the development of national educational systems.[4] Although one of the formally stated goals of PISA is to create an internationally comparative evidence base for educational policy development and implementation (Wiseman, 2013, p.304), Waldow (2009) recognized that headline news about PISA is often more about "shock"[5] over the assessment results than what the assessment information contributes to discussions about long-term educational reform and improvement.

Theoretical and empirical researches (see Štremfel, 2013) show that participating countries become especially attentive to the PISA results when they perform below international (OECD, EU) average. That effect was experienced also in Slovenia. When the PISA 2009 reading literacy results were published and for the first time since Slovenia had been participating in international comparative assessment studies, it showed that Slovenian students perform below international (OECD, EU) average, the perception of the Slovenian educational system as a successful system

4 For more theoretical insight about the role the evaluation plays in the development of public policies see Kristec Lipicer (2009).

5 Phillips and Ochs (2003) explain that education policy shock happens when there is a deviation from the norm, often involving mediocre or low performance (i.e. below expectations).

was marred at the level of experts, policy makers, practitioners and general public (Interviews by author, 2012). PISA 2012 results confirmed the underperformance of Slovenian students in reading literacy and emphasized the need for improvement of the performance of Slovenian students in that domain.

The *aim of the article* is through the understanding of PISA as transnational policy making, using the Slovenian PISA 2012 results, is to show how the policy problem of below average results is identified by participating member states and to illustrate how the policy solutions for the improvement of students' performance in PISA could be found. In order to illustrate the policy framework of improving PISA results, the article as a case study takes into consideration PISA reading literacy results (the domain in which Slovenia perform below OECD and EU average) and students performance at the Proficiency level 2 (the level which Slovenia together with other EU member states chose for defining a common benchmark "to reduce the percentage of low-achieving students to 15% by 2020").[6]

A research question the article addresses is "How to find a way towards improving Slovenia's PISA results according to the concept of transnational policy making and policy learning theory?"

The article addresses the research question in the framework of policy analysis studies. The concept of transnational policy making (in terms of governance of problems and transnational policy promotion) and theory of policy learning (in terms of lesson-drawing) are employed in order to provide an in depth insight into the process of defining and solving policy problems in the contemporary educational policies. Theoretical dispositions are further elaborated on in the case of Slovenian PISA 2012 results in reading literacy and trends in other participating EU member states from 2000 onwards. The empirical data for the case study were gathered by the analysis of the OECD and EU official reports, as well as an analysis of the respective Slovenian legislation and strategic documents. In order to provide an additional understanding of the reception of transnational policy making at the national level, the data gathered by interviews with Slovenian and EU representatives (policy makers, researchers, practitioners) from 2008 to 2012[7] and the results of the survey about the reception of

6 By taking into consideration the policy approaches for improving the PISA results, the article does not take into consideration the more substantive and pedagogical approaches for improving PISA results.

7 Data garhered through semi-structured interviews present an additional source of information and were used only to clarify those open issues that we were unable to identify from our analyses of official documents.

the EU and international agenda in Slovenian educational space conducted in 2012 (Štremfel, 2013) are used.

The article proceeds as follows. In the introduction, the topic and its research framework (question, methods) are explained. In the first section, the article provides insight into how policy problems are constructed in contemporary society with a special emphasis on educational policies through the lenses of the concepts of transnational policy making and new modes of EU governance. In the second section, the possibilities of finding a policy solution for the perceived policy problem are provided using the framework of policy learning theory. Here the article points out two alternative understandings of PISA policy orientation (international policy promotion in the framework of OECD recommendations and lesson-drawing from other participating member states). In the third section, the article, with the help of the case study of Slovenian PISA 2012 reading literacy results elaborates on the difficulties of finding policy solution and improving PISA reading literacy results in Slovenia. In the conclusion, the article summarizes the key findings, which could be taken into consideration by leading the way in order to improve Slovenian PISA results on the basis of lesson-drawing from other successful EU member states.

Identification of National Policy problem(s) Using PISA Results

"Policy problems are those social problems that can be resolved and are being resolved by the state by means of instruments and mechanisms at its disposal" (Fink Hafner, 2002, p. 105). In its widest sense, a policy problem is understood as a deviation between the present situation and a more desired future situation. Processing a problem is usually understood in the sense of solving it. It means that people start thinking about the means of connecting or bridging the gap between what is and what should be. Identification of a policy problem is therefore an important dimension of problem processing. Governing depends on identifying situations as problematic, acknowledging the expertise in connection with these problems and discovering governing technologies that are considered to be a suitable response (Colebatch 2006, p. 313).

From the perspective of social constructivists, the formulation of (mostly transnational) policies turns into the governance of problems. Policy-making actors are present in different spaces and at different times and they differ in terms of their experiences, values, norms and beliefs. Common cooperation is only possible if they succeed in forming a common understanding regarding the necessity of cooperation (Paster, 2005;

Bernhard, 2011). The essential process in relation to this is a joint identification of the problem, which is a prerequisite for cooperation (Hoppe, 2011, p. 50). Governance, as transnational problem solving, takes place when a group of countries recognise a common policy problem and unite their efforts in making plans for its resolution, which is evident from a jointly developed policy model. Governance based on transnational networks in the field of education could not be considered as a national one, as international comparative data construct policy problems and develop policy solutions beyond and between levels (Nóvoa and Yariv-Mashal, 2003; Ozga et al., 2011). Together with the new conception of education (where the main emphasis is on student achievement), the development of new policy instruments (international comparative assessment studies, international comparative achievement scales, benchmarks) guarantees the capacity of governance of the OECD and the European Commission, not only by means of monitoring and assessing national education systems, but mostly by constructing specific policy problems and thus encouraging special assumptions and an understanding of policy learning. Grek (2010) argues that constructing a policy problem is necessary for establishing new modes of governance on the basis of more and more new data, standards and new policy solutions. According to the new modes of EU governance, member states, when they perform below average in PISA comparative achievement scale are faced with triple pressure:

a) Performing below international (OECD, EU) average

Comparisons based on PISA should not be viewed merely as a method, but also as a policy and mode of governance *(governance by comparisons).* Comparisons (commonly shown in international comparative achievement scales) result in definitions of good and bad education systems, legitimise political actions and thus create a new mode of governance. They mostly encompass a rationalistic approach to policy making, wherein (assessed) participants are implicitly under pressure to arrive as close as possible at what is considered 'the best' in accordance with special criteria within a certain context of comparisons. In this regard, the leading assumption is that the most efficient (rationalist approach) and the most suitable (constructivist approach) decisions are adopted on the basis of objective data (March and Olsen, 1989). This objective data, which PISA produces, guarantees the comparability of educational systems and enable member states to identify and eliminate the shortcomings of their educational systems on the basis of mutual comparisons. According to Šenberga (2005, p.15), international comparisons exert positive pressure on national political actors, thereby resulting in policy improvements at the national level.

The meaning being below average in PISA international comparative achievement scale for the prosperity of the participating nation is well elaborated in OECD reports. OECD (2010b, p. 157) argues that "Evidence of the importance of reading literacy for the success of individuals, economies and societies has never been stronger. Past experiences suggest that there are enormous economic gains to be had by OECD countries that can improve the cognitive skills of their populations". The idea that the performance of member states in PISA is an indicator of their further economic development draws attention to PISA results across the world and exerts pressures on participating member states to perform well (that is above average) and therefore ensures their international economic competitiveness and the well-being of their nations.

b) Non-attaining of EU benchmark and common goals

The underlying logic of the concept of governance is that society needs mechanisms for defining common problems, establishing collective goals in order to address and solve these problems, and developing and implementing policy instruments by means of which the goals (outputs) will be achieved (Pierre and Peters, 2000). Grek (2009) believes that within *output-oriented governance*, data and its management play a key role. Data enables governance through goal setting, whereby participant output is directed towards achieving goals. Upon publishing, this data serves as the instruments of encouragement and judgement of participants in terms of their output. It thus simultaneously controls the autonomy of the actors operating within the context in relation to how they will achieve their goals. This is a system of discipline based on the judgement and classification of participants in achieving (jointly defined) goals.

Grek (2009) argue that one of the most visible examples of output-oriented governance is common EU cooperation in the field of education. In order that their educational systems could importantly contribute to the development of EU smart, sustainable and inclusive growth as defined in EU 2020 strategy, EU member states agreed on common educational goals, benchmarks and indicators, which they follow and monitor together. PISA data is used for one of the benchmarks, which states that by 2020, the share of 15-year-olds with a low achievement in reading, mathematics and science should be less than 15%.

Since a benchmark has been commonly agreed, member states feel responsible to effectively contribute to its attainment. The member states' responsibility is strengthen also by the publications of the European Commission (see for example European Commission, 2013), which by analysing the progress towards attaining particular common goal at the EU lev-

el, points out the member states, which the most and the less successfully contribute to its attainment. Authors (e.g. Alexiadou, 2007) argue that these publications present a subtle pressure on member states on the basis of "naming and shaming" and encourage member states to improve their PISA results and consequently effectively contribute to commonly agreed benchmarks and goals.

c) Non-attaining of national goals

Empirical study (Štremfel, 2013) revealed that in Slovenia, the international comparative assessment studies (including PISA) are regarded as an objective indicator of the knowledge of Slovenian students and that they allow identification of policy problems when it comes to Slovenia's below-average achievements in comparison with the international (EU and OECD) average. The importance of Slovenian performance in these studies is highlighted also in the White Paper on Education (2011, p. 25), where it is stated: "At the state level we need to state and map out a clear path towards the goal, that performance of Slovenian students in international comparative assessment studies are at the top, that mean at least in the upper third of the students' achievement of the developed countries".

The international comparative assessment studies are therefore understood as an instrument of external evaluation of the national educational system. Public policy evaluation is generally defined as any assessment of the public policy effects and provides information, whether the objectives of the public policies are being achieved (Dye, 1995, p. 321). When the results of the evaluation show that public policy does not successfully follow its objectives, it calls for the abolition of public policy or its improvement.

In this section, it is presented how non-attainment of the national educational objectives (related also to the (below average) performance of students in international comparative assessment studies) is theoretically perceived as a policy problem. In the next section, the article presents two distinct theoretical ways of resolving these policy problems through the lenses of transnational policy making.

PISA Policy Impact: From International Policy Promotion to Lesson-drawing

Authors (e.g. Meyer and Benavot, 2013) argue that PISA does not allow only the identification of policy problems, but on the basis of its results also policy solutions can be provided. That is usually understood in terms of PISA strong policy orientation and its policy impact on member states (Grek, 2010). We consider that PISA policy impact in terms of problem

solving can be understood from two distinct types of policy learning: international policy promotion and lesson-drawing (see Holzinger and Knill, 2006).

a) Policy impact as international policy promotion

Cross-national policy learning is stimulated by the construction of international comparative achievement scales ranking national policies in terms of performance to previously agreed criteria (Grek, 2009). In constantly searching for new policy ideas, disseminating best practice and evaluating national policy performance, international institutions (also OECD) function as mediators of cross-national policy learning, urging national governments to adopt successful policy models (Kern et al., 2000, p. 10 in Holzinger and Knill, 2006, p. 22). Since it is believed that international institutions promote the spread of distinct policy approaches they consider particularly promising, this process is understood as international policy promotion. Countries that deviate from recommended policy models or rank low in international comparative achievement scale face pressure to legitimate their policy approaches in light of "international scrutiny" and are motivated to adopt these certain policy approaches because of legitimacy pressures of the international institutions (Holzinger and Knill, 2006, p. 22).

Carvalho (2012, p. 173) argues that having in mind the concept of a public policy instrument, one may say that PISA is driven by a specific "problematisation" of the role of education in contemporary times and by a specific model for the regulation of the educational sector. With its international comparative achievement scales publicising which countries are progressing in the right direction and which are falling further behind with respect to student achievement, PISA steers participating nations towards a particular model of curricular and structural reform (Takayama, 2012, p. 148).

Recommendations resulting from expert discourse are based on the strategy of comparison and attempt to impose similar answers for different national contexts. In different countries, the OECD recommendations have been accepted as valid among policy makers and stakeholders on the basis of the authoritative characteristic of knowledge included in these reports (Grek, 2010, p. 398). An important factor of the readiness to accept these recommendations is uncertainty with regards to how to improve their results in international comparative achievement scales (Nóvoa and Yariv-Mashal, 2003; Grek, 2010). Under this approach, then, PISA is seen as a way of gradually solving national problems by moving problem solving capacity from the national to the supranational level (see also Alexiadou, 2014, p. 128).

Bieber and Martens (2011) explain, that OECD from the PISA results draw recommendations for policy-making by concentrating on factors that are positively correlated with student performance in PISA, though without claiming a causal relationship. These recommendations (included in international thematic reports or specific country reviews) range from rather implicit to very explicit statements. For example, OECD Economic Review for Slovenia, which refers also to PISA results, provides the recommendations considering the efficiency of Slovenian basic education. In the Review (OECD, 2011, p. 1), it is stated that "saving could be gained by enhancing spending efficiency in early childhood and basic education, which are plagued by high costs due to low pupil-teacher ratios, small class sizes and high numbers of non-teaching staff. Merging schools and extending catchment areas, while taking into account other socio-economic considerations, could bring significant efficiency gains". Therefore, the OECD Economic Review, by proposing very concrete and economic oriented measures, which does not take into consideration the particularities of the Slovenian national context (Educational Research Institute, 2011) could be seen as international policy promotion.

Although international policy promotion in the situation of uncertainty about how to improve PISA results could be an attractive idea for participating member states, even more so when PISA is understood as an objective and neutral evaluation of a national education system (Interviews, 2012). It is worth mentioning that such international policy promotion erodes the traditional idea of member states sovereignty over their national educational systems (e.g. Walkenhorst, 2008; Zgaga, 2011). On one hand, some authors are concerned about international policy promotion and see it as a portfolio of best practices imposed to national governments by global actors (e.g. OECD, EU). On the other hand, some authors (e.g. Steiner-Khamsi, 2012, p. 3) argue that travelling reforms supposedly represent best practices or international standards that have been transferred successfully from one country to another and regard policy making as a rational undertaking, and view policy learning as examples of lesson-drawing, thus one of the more desirable outcomes of evidence-based policy making.

b) Policy impact as lesson-drawing

Lesson-drawing[8] is seen as a pragmatic tool for identifying and transferring "best practices" from one context to another with the goal of solv-

8 A policy lesson according to Rose (1993, p. 27) is "a detailed cause-and-effect description of a set of actions that government can consider in the light of experience elsewhere, including a prospective evaluation of whether what is done elsewhere could someday become effective here".

ing problems and improving educational systems in different national settings (Rose, 1991; 1993). In the case of lesson-drawing, an individual country searches for foreign examples and *decides on its own* to what extent and in which way it will "learn from others" when modifying, improving or making new national policy. A particular country therefore voluntary decides on its own from which country it will learn from and to what extent, as well as how, it will monitor any (new or amended) policy adjustments (including its implementation), and to whom – if anybody at all – it will report its success to (Fink-Hafner et al., 2010, p. 19).

In the lesson-drawing the decisions are based on searching for the means to pursue valued goals in a systematic and comprehensive manner, reviewing policy in the light of past experience and other available information to make adjustments where necessary (James and Lodge, 2003, p. 181). The presumption is that actors work in rational accounts. The question of "how to improve", guides specific mechanisms for improving, including sources and ways to analyse evidence (James and Lodge, 2003, p. 190). Lesson-drawing (when used for resolving identified policy problem, improving national educational policy and consequently improving PISA results) therefore requires serious scientific investigation.[9]

According to Rose (2002) lesson-drawing should be implemented very carefully by following ten steps: (1) Diagnosing your problem; (2) Examining where to look for a lesson; (3) Investigating how a programme works there; (4) Abstracting a cause-and-effect model for export; (5) Designing a lesson; (6) Deciding should it be imported?; (7) Identifying resource requirements and constraints; (8) Exploring the problem of context; (9) Bounding speculation through prospective evaluation; (10) Identifying foreign countries as positive or negative symbols.

After presenting two theoretical insights in PISA policy orientation and policy impact (international policy promotion and lesson-drawing), it is interesting to see how OECD itself understands PISA policy orientation. OECD (2003, p. 16) states: "Key features driving the development of PISA have been: its policy orientation, with design and reporting methods determined by the need of governments to draw policy lessons". As seen the OECD definition does not involve any specific type of policy learning and therefore (at least officially) leaves the space for employing different types of policy learning from PISA results open.

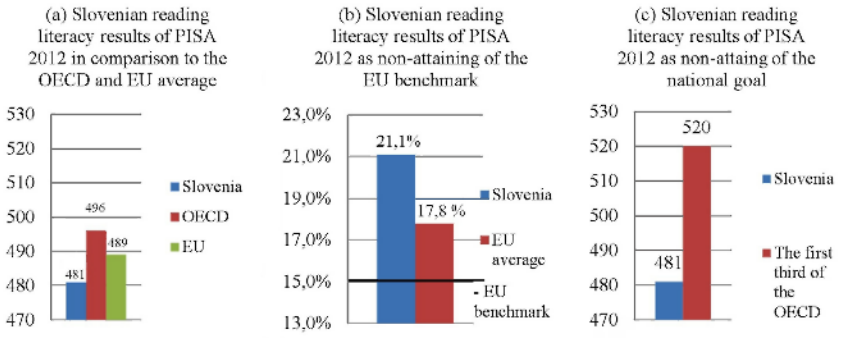
9 Phillips (2013, p. 299) argues that in addition to (a) serious scientific investigation there could be also other motives for lesson-drawing: (b) popular conceptions of the superiority of other approaches to the educational questions; (c) politically motivated endeavours to seek reform of provision by identifying a clear contrast with the situation elsewhere; (d) distortion (exaggeration), whether or not deliberate, of evidence from abroad to highlight perceived deficiencies at home.

Taking into consideration the importance of preserving the sovereignty of national states over the development of their educational systems (e.g. Walkenhorst, 2008; Zgaga, 2011) and the Slovenian experience with the de-contextualized OECD recommendations (2011) to its educational system, we consider that it is important for participating member states (also Slovenia) to find their own policy solution to identified policy problem (below-average PISA results). In the next section, the article therefore provides a more detail empirical insight in the lesson-drawing as a promising strategy for improving below-average PISA results using the case study of Slovenia.

PISA Policy Problem in Slovenia and a Way Towards its Solution

When a state identifies policy problem according to its below-average performance in PISA and decides to solve it by the lesson-drawing (drawing lessons from other (successful) participating states), there are some theoretical dispositions developed which can assist and guide individual states towards that comprehensive process. In this section, the article tries to provide some empirical insight into first two of the Rose (2002) ten steps for learning from abroad, presented in the previous section.

Figure 1: Slovenian reading literacy results of PISA 2012 as a policy problem



Source: OECD (2013a); European Commission (2013).

The first step of lesson-drawing, according to Rose (2002), presents the *identification of the national policy problem*. Rose (ibid) argues that when political dissatisfaction is high, and especially if it is unexpected, there is often confusion about what exactly the problem is. Figure 1 shows, why Slovenian reading literacy results of PISA 2012 could be understood as a policy problem according to the three perspectives presented in the

first section of the article ((a) performing below international (OECD, EU) average; (b) non-attaining of EU benchmark and common goals; (c) non-attaining of national goal)).

Figure (a) shows that mean score of Slovenian students in PISA 2012 reading literacy was 481, while the OECD average was 496 and the EU average 489. Figure (b) shows that 21,1 % of Slovenian students could be considered as low achieving students according to PISA 2012 reading literacy results (e.g. not attaining the second (basic) level of reading literacy). That means that Slovenia does not successfully follow the EU benchmark of 15% of low achievers in PISA by 2020. Figure (b) also shows that percentage of low-achieving students in reading literacy in Slovenia (21,1 %) is higher than on the average in the EU member states (17,8 %). Figure (c) shows that taking into consideration PISA 2012 reading literacy results, the Slovenian long-term goal "to perform in the first third of the most developed countries"[10] was not reached. The average main score of the first third OECD states was 520, the main score of Slovenia was 481. Results presented in figures 1 (a), (b), (c) therefore show that according to all three criteria performances of Slovenian students in PISA 2012, reading literacy could be perceived as a policy problem.

It should be noted that according to all three perspectives, the policy problem was recognized with reference to the external measures. Even the national goal was stated in terms of ranking in international comparative achievement scale and does not provide a detailed insight in more substantive national goals and priorities. According to Rose (2002, p. 6) stating that "there is no point in looking abroad for a remedy if you do not know what the problem is at home", Slovenia has not realized the first Rose step of successful lesson-drawing yet. In order to empirically explain how non defined national goals and priorities can hinder the further process of lesson-drawing according to Rose (2002), at this point of the article we move to his second step of lesson-drawing that is to the question "Where to look for a lesson"?

Rose (1993, p. ix-x) argues that lesson-drawing occurs across time and space and is both positive, leading to prescriptions about what ought to be done, and negative, in terms what not to emulate. Although there are some suggestions about the usefulness of concentrating on the failure of other member states, (see Radaelli, 2004), authors (e.g. Hemerijck and Visser, 2001) it is argued that it is more promising to look for lessons from

10 Although in the Slovenian White Paper on Education it is not exactly defined, what is considered under »the most developed countries«, we took into consideration the results of the OECD member states. OECD is often called »the club of world's most advanced countries« (OECD, 2014).

Table 1: Identification of the most successful EU member states in following the EU benchmark

PISA cycle/ member state	2000	2003	2006	2009	2012	Trend (2000–2012)
Austria	19.9%	21.7%	21.5%	/	19.5%	−0.4%
Belgium	19.0%	17.9%	19.4%	17.7%	16.1%	−2.9%
Bulgaria	42.5%	/	51.1%	41.0%	39.4%	−3.9%
Croatia	/	/	21.5%	22.5%	18.7%	/
Czech Republic	17.5%	19.4%	24.8%	23.1%	16.9%	−0.6%
Cyprus	/	/	/	/	32.8%	/
Denmark	17.9%	16.5%	16.1%	15.2%	14.6%	−3.3%
Estonia	/	/	13.6%	13.3%	9.1%	/
Finland	7.1%	5.7%	4.8%	8.1%	11.3%	+4.3%
France	15.2%	17.5%	21.7%	19.8%	18.9%	+3.7%
Germany	**22.6%**	**22.3%**	**20.0%**	**18.5%**	**14.5%**	**−8.1%**
Greece	24.4%	25.3%	27.7%	21.3%	22.6%	−1.8%
Hungary	22.7%	20.5%	27.6%	17.6%	19.7%	−3.0%
Ireland	11.1%	11.2%	12.1%	17.2%	9.6%	−1.4%
Italy	18.9%	23.9%	26.4%	21.1%	19.5%	+0.6%
Latvia	36.1%	18.1%	21.2%	17.6%	17.0%	−15.1%
Lithuania	/	/	25.7%	24.3%	21.2%	/
Luxemburg	/	22.7%	22.9%	26.0%	22.2%	/
Netherlands	/	11.5%	15.1%	14.3%	14.0%	/
Poland	**23.2%**	**16.8%**	**16.2%**	**15.0%**	**10.6%**	**−12.6%**
Portugal	26.6%	21.9%	24.9%	17.6%	18.8%	−7.8%
Romania	41.3%	/	53.5%	44.4%	37.3%	−4.0%
Slovakia	/	24.9%	27.8%	22.2%	28.2%	/
Slovenia	/	/	16.5%	21.1%	21.1%	/
Spain	16.3%	21.1%	25.7%	19.6%	18.3%	+2.0%
Sweden	12.6%	13.3%	15.3%	17.4%	22.7%	+10.1%
United Kingdom	/	/	19.0%	18.4%	16.6%	/

Source: OECD (2013a).

those who succeed. However, which educational system could be considered as successful according to the PISA results? Although we agree that there is no one way to answer this, in this article, we adopt the OECD (2010a, p. 14) understanding of successful states as not just top-scoring,

but especially those ones which are rapidly improving from the first PISA cycle in 2000 onwards. OECD (2010a, p. 13) explains:

"The design of PISA does not just allow for a comparison of the relative standing of countries in terms of their learning outcomes; it also enables each country to monitor changes in those outcomes over time. Such changes indicate how successful education systems have been in developing the knowledge and skills of 15-year-olds. All countries seeking to improve their results can draw encouragement – and learn lessons – from those that have succeeded in doing so in a relatively short period of time."

Table 1 shows the trends of the EU member states' PISA reading literacy performance since 2000 in order to identify those member states, which were the most successful in improving the results of their low-achieving students and the most successfully follow the EU benchmark "to reduce percentage of low-achievers in PISA to 15% by 2020".[11]

The Table 1 shows that among 28 EU member states, 18 of them have been participating in all PISA cycles (2000-2012). For these member states trends in the percentage of low-achieving students are presented. It shows that 13 of the member states succeed in reducing the percentage of their low achieving students, and in 5 of them, these percentages from 2000 to 2012 increase. The table also shows that the most successful EU member states in reducing the percentage of their low-achieving PISA students in reading literacy are Poland (12,6%), and Germany (8,1%).

OECD (2010a) claims that success of such a diverse group of countries in raising the level of their students' performance in reading indicates that improvement is possible regardless of a country's context and where it starts out from. Similarly, European Commission (2013) recognizing that the EU as a whole is lagging behind in its challenge to reduce the share of low achievers in reading, points out that this trend does, however, disguise large differences found between and within EU member states. By indicating the concrete member states and their improvement, the European Commission does not only exert the pressure on some member states on the basis of "naming and shaming" but also indicate the countries, from which the lessons could be drawn. The European Commission (2013, p. 5) states: "The reasons why some member states succeeded in significantly reducing the share of low achievers may serve as an inspiration for other countries that are struggling to overcome similar challenges or even face a deteriorating situation."

11 Although OECD identifies trends in results of participating countries on a special methodology (see OECD, 2013a) which measure trends only between the cycles when the math was a main testing domain, we have present trends from the 2000 onwards which is also the established practice of the EU.

Table 2: Overview of the performance of the most successful member in following the EU benchmark by different indicators

Indicator / member state	Slovenia	Poland	Germany
Percentage of low achievers (2000)	/	23.2%	22.6%
Percentage of low achievers (2012)	27.1%	10.6%	14.5%
Percentage of low achievers (difference 2012-2000)	/	12.6%	8.1%
Percentage of high achievers (2000)	/	5.9%	8.8%
Percentage of high achievers (2012)	5.0%	10%	8.9%
Percentage of high achievers (2012-2000)		4.1%	0.1%
Gap between 90th and 10th percentiles (2000)	/	260 points	284 points
Gap between 90th and 10th percentiles (2012)	236 points	222 points	237 points
Change in gap between 90th and 10th percentiles (2012-2000)	/	38 points	47 points
Proportion of total variation explained by between-school variance (2000)	/	62%	59%
Proportion of total variation explained by between-school variance (2009)	41.8%	65.4%	67.2%
Change in proportion of total variation explained by between-school variance (2009-2000)	/	3.4%	8.2%
Relationship between reading performance and the PISA index of economic, social and cultural status (ESCS) (2000)	/	40 points	52 points
Relationship between reading performance and the PISA index of economic, social and cultural status (ESCS) (2009)	/	39 points	44 points
Change in relationship between reading performance and the PISA index of economic, social and cultural status (ESCS) (2009-2000)		1 point	8 points
Difference in performance between native students and students with immigrant background (2000)	/	/	84 points
Difference in performance between native students and students with immigrant background (2009)	/	/	56 points
Change in difference in performance between native students and students with immigrant background (2009-2000)	/	/	-28 points

Source: OECD (2010a; 2013a; 2013b).

Since some authors (see Štremfel et al., 2014) argue that one indicator cannot provide enough insight in the functioning of the individual educational system, Table 2 shows how in EU member states, which

states succeed the most in following the EU benchmark (reduce a number of low-achievers in reading literacy), the which trends in other indicators have changed.

Table 2 shows that Poland and Germany, which succeed the most in following EU benchmark (reducing the percentage of low-achievers), were not as successful in other selected indicators. Estimating which of them would be the most appropriate to learn from in order to improve Slovenian PISA results and successfully follow the EU benchmark is therefore a comprehensive task. The review of trends in different indicators shown in Table 2, first of all requires that a learning country (Slovenia) define concrete goal about which set of indicators it would like to improve upon. One single benchmark (defined at the EU level) is too broad and cannot provide that focus and learning the state should find itself. Even OECD (2010a, p.4) recognized that "PISA results suggest that the countries that improved the most, or that are among the top performers, are those that establish clear, ambitious policy goals (...)."

Conclusions

If a new mode of governance in the EU is viewed as governing, steering and supervising actors (Kooiman, 2003, p. 3), for them to participate in collective policy problem solving and thus achieve the pursued goals jointly (Pierre and Peters, 2000), the highlighted lack of clarity of educational goals both at the supranational and the national level: (a) opens up room for political manipulation of international organisations (Borrás and Radaelli, 2011) or (b) present a huge obstacle on the way of improving the results on the basis of lesson-drawing. The wideness and openness of goals (and consequently their lacking clarity) allows the development of legitimate, reasonable and good policies and the (imaginary) common good in the context of social learning (Borrás and Conzelmann, 2007) and therefore pursuing a specific not necessary evidence-based educational model.

With apparent PISA neutrality EU and OECD steers the member states towards achieving specific educational goals. The EU benchmark (reducing a number of low achievers to 15% by 2020) facilitates assessments and comparisons of member states' achievements *(output-oriented governance and governance by comparison)* in pursuing the common EU goals. PISA comparative achievement scale thus exerts dual pressure on the EU member states. The primary pressure to perform well is related to securing the international competitiveness of the state. The secondary pressure to perform well is related to avoiding the blaming and shaming by the European Commission and by other member states for not attaining com-

mon agreed goals (Alexiadou, 2007; Ioannidou, 2007). Once a member state perceives a policy problem (related to lack of economic competitiveness) following its ranking on PISA achievement scale, the best models for solving the problems in question *(governance of problems)* have commonly already been developed at the OECD level. In the article, this is shown using the example of OECD Economic Review for Slovenia (2011). In the case that member states follow these recommendations, the presented dynamics facilitates the deepening of the OECD cooperation in the field of education towards what is preferred by the OECD *(international policy promotion)*, while the member states have over the past few years – in the circumstances of the economic crisis – been following the OECD more so than before, aiming to maintain their competitiveness within the knowledge-based economy (also see Tsarouhas, 2009).

However, it is also necessary to be aware of the fact that actors have different sources for a critical appraisal of the knowledge provided by international comparative assessment studies and an effective use of that knowledge for development of their national educational systems. In such a context, deep and careful reflection about the nature of knowledge and its mobilisation within public policy is essential. This raises a question of whether the use of (international) comparisons as a mode of governance has not resulted in excessive legitimacy of knowledge they produce and whether it is time for actions towards a diversity of knowledge types, communicated by means of knowledge-based governance tools (Delvaux and Mangez, 2010).

The main implications of understanding PISA as transnational problem solving would therefore be that the expert knowledge, which the PISA and other international comparative assessment studies provide, should be used at the national level in accordance with neopositivist and critically rational means of "speaking truth to power" and not in accordance with the interpretative and neopragmatic means of "making sense together" (Hoppe 2011, p. 55).[12] In this author's opinion, the role of national experts is to assess what data (from PISA and international comparative assessment studies) and proposals for solving the identified policy problems are to be taken as legitimate and definite in implementing the changes and improvements in the national system (Wiseman 2010, p. 9). That was already recognized, when the OECD Economic Review for Slovenia

12 Experts and the expert knowledge would thus be used in an instrumental sense of making the right decisions and not for the advocacy of political decisions and the ideology of (supranational and national) political actors (Stone, 2001; Jones, 2009; Nasschi and Demszky, 2011). After all, in Slovenia some in-depth critical deliberations regarding the (non-)use of expert knowledge in education policy making have already been undertaken as well (Gaber, 2007; Kodelja, 2007) and might be worth reconsidering.

was published and national educational experts warned about some of its misleading conclusions.

The appropriate use of expert data, which can be acquired from PISA and other international comparative assessment studies, can facilitate the preservation of distinct national characteristics and the quality of the education system and make thoughtless acceptance of international policy promotion of certain educational models much more unlikely (Grek, 2008). In order to preserve the sovereignty of national state over its educational system, the article therefore suggests that instead of an uncritical reception of the international promotion of certain educational model (OECD, 2010c; Hanuschek and Woessmann, 2011), the more promising alternative for improving PISA results is lesson-drawing.

If we understand the policy problem as the deviation between the actual current situation and the desired future situation, more emphasis in Slovenia should be put on the concretization of the desired future situation in terms of concretization of the national educational goals. The argument is that the EU benchmark (to reduce number of low-achievers to 15% by 2020) and Slovenian strategic goal (to perform in the first third of the developed countries in international comparative assessment studies) are too broad to identify concrete policy problem, to provide a solid base for national educational reform and to target specific policy measures for improving PISA results and pursuing these goals. In addition, Table 2 shows that monitoring one single indicator (reducing number of low-achievers to 15% by 2020) is not sufficient for in depth understanding of long-term performance of successful member states, which is necessary in order to draw lessons from them and to find their own way of improving PISA results. One single benchmark / indicator therefore cannot provide an in depth insight into which policy measures EU member states should focus on, which measures have succeeded the most in reducing the number of low-achievers in their educational system (Germany, Poland), which have been employed and how these policy measures have impacted other indicators, as well as which are important for ensuring equity and quality in their educational system. Since the lesson-drawing is a timely and expensive process (Rose, 2002), it is even more important that states do not make a mistake already in the first two steps of learning from abroad (identifying of policy problem and finding a state, from which they will learn from). If we took into consideration the theoretical dispositions (of *international policy promotion* and *lesson-drawing*) presented in the article, it seems that the main policy lesson for Slovenia from PISA results is that clear goals should be stated and then followed to more overreaching goals – not only to perform in the first third of the most developed

participating countries in international comparative achievement scales, but more importantly to be aware (on the all levels of the educational systems), which educational goals we are following in order to contribute to the welfare of the nation. PISA results should therefore be seen as the external mirror for finding and monitoring member states own ways of improving educational results and not a goal in itself.

References

Alexiadou, N. (2007) The Europeanisation of education policy – changing governance and »new modes of coordination«. *Research in Comparative and International Education.* 2 (2), pp. 102-116.

Alexiadou, N. (2014) Policy learning and Europeanisation in education: The governance of a field and the transfer of knowledge. In: Nordin, A. & Sundeberg, D. (eds.) *Transnational Policy Flows in European Education: the making and governing of knowledge in the education policy field.* Oxford: Symposium Books.

Berhard, S. (2011) Beyond Constructivism: The Political Sociology of an EU Policy Field. *International Political Sociology.* 5 (4), pp. 426-445.

Bieber, T., and Martens, K. (2011) The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US. *European Journal of Education.* 46 (1), pp. 101-116.

Borrás, S., and Conzelmann, T. (2007) Democracy, Legitimacy and Soft Modes of Governance in the EU: The Empirical Turn. *Journal of European Integration.* 29 (5), pp. 531-548.

Borrás, S., and Radaelli, C. (2011) The Politics of Governance Architectures: Creation, Change and Effects of the EU Lisbon Strategy. *Journal of European Public Policy.* 18 (4), pp. 463-484.

Carvalho, L. M. (2012) The Fabrications and Travels of a Knowledge – Policy Instrument. *European Educational Research Journal.* 11 (2), pp.172–188.

Colebatch, K. H. (2006) What work makes policy? *Policy Sciences.* 39 (4), pp. 309–321.

Council of the EU (2009). Council conclusions of 12 May 2009 on a strategic framework for European cooperation in education and training ('ET 2020'). Brussels: Council of the European Union.

Delvaux, B., and Mangez, E. (2010) Knowledge and policy: a new deal. Leuven: University of Leuven: KNOWandPOL Research Report.

Dye, T. R. (1995) *Understanding public policy.* Englewood Cliffs, New Jersey: Prentice Hall.

Educational Research Institute (2011) Kritična tiskovna konferenca ob poročilu OECD o slovenskem šolstvu. / Critical press conference about

the OECD Economic Review for Slovenia. Ljubljana, 16th March 2011.

European Commission (2013) *PISA 2012: EU performance and first inferences regarding education and training: policies in Europe.* Brussels: European Commission.

Fink-Hafner, D. (2002) Analiza politik – akterji, modeli in načrtovanje politike skupnosti. In Jelovac, D. (ed.). *Jadranje po nemirnih vodah menedžmenta nevladnih organizacij. / Sailing troubled waters of management of non-government organizations.* Ljubljana: Radio Študent: Študentska organizacija Univerze v Kopru: Visoka šola za management. pp. 105-123.

Fink-Hafner, D., Lajh, D., and Deželan, T. (2010) The Open Method of Coordination in the Global Context of Policy Cooperation. In: Fink-Hafner, D. (ed.) *The Open Method of Coordination.* Ljubljana: Faculty of Social Sciences.

Gaber, S. (2007) Spoprijem za hegemonijo ali vaje iz praktične teorije? *Sodobna pedagogika* [Contemporary Pedagogy]. 58 (2), pp. 62–81.

Grek, S. (2008) From symbols to numbers: the shifting technologies of education governance in Europe. *European Educational Research Journal.* 7 (2), pp. 208–218.

Grek, S. (2009) The role of the OECD as an agent of Europeanization: problematisation and change in education governance in Europe. Paper presented at Edinburgh's Europa Institute Seminar: Practicing EU Government: Problematisation, 6. May, Edinburgh, Great Britain.

Grek, S. (2010) International Organisations and the Shared Construction of Policy »Problems«: problematisation and change in education governance in Europe. *European Educational Research Journal.* 9 (3), pp. 396-406.

Hanuschek, E., A., and L. Woessman, L. (2011) The Cost of Low Educational Achievement in the European Union. EENEE Policy Brief.

Hemerijck, A., and Visser, J. (2001) Learning and Mimicking: How European Welfare States Reform. Amsterdam: University of Amsterdam, Working Paper.

Holzinger, K., and Knill, C. (2006) Causes and conditions of cross-national policy convergence. In: Knill, C. (ed.). *Cross-national Policy Convergence: Concepts, Causes and Empirical Findings.* London: Routledge.

Hopmann, T. S., Brinek, G., and Retz, M. (2007) *PISA 'According to PISA. Does PISA Keep What it Promises.* New Brunswick in London: Transaction Publishers.

Hoppe, R. (2011) *The Governance of Problems: Puzzling, Powering and Participation*. Bristol: The Policy Press.

Ioannidou, A. (2007) A Comparative Analysis of New Governance Instruments in the Transnational Educational Space: a shift to knowledge-based instruments? *European Educational Research Journal*. 6 (4), pp. 336-347.

Interviews by author (2012) Interviews conducted with public officials, experts and practitioners in the field of education policy. In: Štremfel, U. (2013). Nova oblika vladavine v Evropski uniji na področju izobraževalnih politik. [New modes of governance in the EU in the field of education policies.], pp. 400. Ljubljana: Faculty of social sciences.

James, O., and Lodge, M. (2003) The Limitations of 'Policy Transfer' and 'Lesson Drawing' for Public Policy Research. *Political Studies Review*. 1 (2), pp. 179-193.

Jones, H. (2009) Policy-making as discourse: a review of recent knowledge-to-policy literature. A Joint Information and Knowledge Management and Overseas Development Institute Emergent Working Paper No. 5.

Kodelja, Z. (2005) Komparativne edukacijske raziskave in šolska politika / Comparative educational research and school policy. *Šolsko polje*. 16 (3-4), pp. 211–226.

Kodelja, Z. (2007) Reforme, stroka in šolske politike. *Sodobna pedagogika* [Contemporary Pedagogy]. 58 (2), pp. 34–48.

Kooiman, J. (2003) *Governing as Governance*. London: Sage.

Kustec Lipicer, S. (2009) *Vrednotenje javnih politik*. Ljubljana: Fakulteta za družbene vede.

Lasswell, H. D. (1951) The Policy Orientation. In: Lerner, D., and Lasswell, H. D. (eds.) *The Policy Sciences. Recent Development in Scope and Method*. Stanford: Stanford University Press.

March, G. J., and Olsen, P. J. (1989) *Rediscovering Institutions. The Organizational Basis of Politics*. New York: The Free Press.

Meyer, H. D. and Benavot, A. (2013) *PISA, Power, and Policy: the emergence of global educational governance*. Oxford: Symposium Books.

Nasschi, A. von der, Hagen-Demszky, A., and Mayr, K. (2007) The Structures of Knowledge and of Knowledge Production. KNOWandPOL Report, Literature review, Part 8.

Novoa, A., and Yaric-Marshal, T. (2003) Comparative Research in Education: A mode of governance or a historical journey? *Comparative Education*. 39 (4), pp. 423-443.

OECD (2003) *Literacy Skills for the World of Tomorrow – Further results from PISA 2000.* Paris: OECD.

OECD (2010a) *PISA 2009. Results, Learning Trends: Changes in Student Performance Since 2000.* Paris: OECD.

OECD (2010b) *PISA 2009 Results: What Students Know and Can Do. Student performance in reading, mathematics and science.* Paris: OECD.

OECD (2010c) *The High Cost of Low Educational Performance. The long-run economic impact of improving PISA outcomes.* Paris: OECD.

OECD (2011) *OECD Economic Surveys Slovenia.* Paris: OECD.

OECD (2013a) *PISA 2012 Results: What Students Know and Can Do. Student performance in mathematics, reading and science.* Paris: OECD.

OECD (2013b) *PISA 2012 Results: Excellence through Equity. Giving every student the chance to succeed.* Paris: OECD.

OECD (2014) OECD members and partners. Available from: http://www.oecd.org/about/membersandpartners/. (Accessed: 24th May 2014).

Ozga, J., Dahler-Larsen, P., Sergerholm, C., and Simola, H. (eds.) (2011) *Fabricating Quality in Education: Data and Governance in Europe.* London: Routledge.

Paster, T. (2005) The New Modes of EU Governance: Combining Rationalism and Constructivism in Explaining Voluntarist Policy Coordination in the EU. *Österreichische Zeitschrift für Politikwissenschaft.* 34 (2), pp. 147-161.

Pereyra, M. A., Kottoff, H. G., and Cowen, R. (eds.) (2011) *PISA under examination. Changing Knowledge, Changing Tests, and Changing Schools.* Rotterdam: Sense Publishers.

Phillips, D., and Ochs, K. (2003) Processes of policy borrowing in education: Some analytical and explanatory devices. *Comparative Education.* 39 (4), pp. 451-461.

Pierre, J., and Peters, G. B. (2000) *Governance, Politics and the State.* New York: St. Martin's Press.

Radaelli, C. (2004) Who learns what? Policy learning and the open method of coordination. Paper presented at Economic and Social Research Council Workshop on the Lisbon Strategy, University of Birmingham, 26. November, Birmingham, Great Britain.

Rose, R. (1991) What is Lesson Drawing? *Journal of Public Policy.* 11 (1), pp. 3-30.

Rose, R. (1993) *Lesson-Drawing in Public Policy: A Guide to Learning across Time and Space.* Chatham: Chatham House Publishers.

Rose, R. (2002) Ten Steps in Learning Lessons from Abroad. EUI Working Paper. Florence: European University Institute.

Scharpf, W. F. (1997) Introduction: The Problem-Solving Capacity of Multi-Level Governance. *Journal of European Public Policy.* 4 (4), pp. 520-538.

Steiner-Khamsi, G. (2012) Understanding Policy Borrowing and Lending. Building Comparative Policy Studies. In: Steiner-Khamsi, G., and Waldow, F. (eds.) *Policy Borrowing and Lending in Education.* London and New York: Routledge.

Stone, D. (2000) Learning Lessons, Policy Transfer and the International Diffusion of Policy Ideas. Warwick: University of Warwick, Centre for the Study of Globalisation and Regionalisation Working Paper, No. 69/01.

Šenbrega, K. (2005) The Open Method of Coordination and Mutual Policy Learning, Opportunities and Constraints. Lund: Lund University, Department of Political Science, University Essay from Lund University.

Štremfel, U. (2013) Nova oblika vladavine v Evropski uniji na področju izobraževalnih politik. / New modes of governance in the EU in the field of education policies. Ljubljana: Faculty of social sciences.

Štremfel, U., Štraus, M. and Lajh, D. (2014) Sprejemanje rezultatov PISA 2009 v Sloveniji skozi prizmo javnopolitičnega učenja. *Šolsko polje.* 25 (1-2), pp. 79-104.

Takayama, K. (2012) Bringing a Political "Bite" to Educational Transfer Studies: Cultural Politics of PISA and the OECD in Japanese Education Reform. In: Steiner-Khamsi, G., and Waldow, F. (eds) *Policy Borrowing and Lending in Education.* London: Routledge.

Tsarouhas, D. (2009) The Open Method of Coordination and integration theory: are there lessons to be learned? Paper presented at 11th European Union Studies Association Biennial Conference, 23.–25. April, Los Angeles, USA.

Waldow, F. (2009) What PISA Did and Did Not Do: Germany after the "PISA-shock". *European Educational Research Journal.* 8 (3), pp. 476-483.

Walkenhorst, H. (2008) Explaining change in EU education policy. *Journal of European Public Policy.* 15 (4), pp. 567-587.

White Paper on Education (2011) Ljubljana: Ministry of Education, Science and Sport.

Wiseman, A. W. (2013) Policy Responses to PISA in Comparative Perspective. In: Dieter Meyer, H., and Benavot, A. (eds.) *PISA, Power and Policy. The emergence of global educational governance.* Oxford: Symposium books.

Zgaga, P. (2011) Education for "a better world": is it still possible? *Education Inquiry.* 2 (2), pp. 331-343.

Looking Back at Five Rounds of PISA: Impacts on Teaching and Learning in Germany

Christine Sälzer and Manfred Prenzel

The 3rd December 2013 was quite a happy day in Germany. When the results of PISA 2012 were released to the public, the picture looked gratifying and the conclusion to be drawn was that Germany did well in educating their youth and preparing young people for their lives after compulsory schooling. The fact that Germany performed well above the OECD average in all three PISA domains was considered good news, although there was still a gap between Germany and the internationally top-performing countries. If PISA 2012 had been the first round, the results probably would have been received less positively. In contrast, with reference to the results of the first round, known as the German "PISA-shock", it becomes quite clear what a distance lies between the student performance in PISA 2000 and in PISA 2012. The above-average results in PISA 2012 can therefore be seen as a common achievement between several players in national education policy, research and practice.

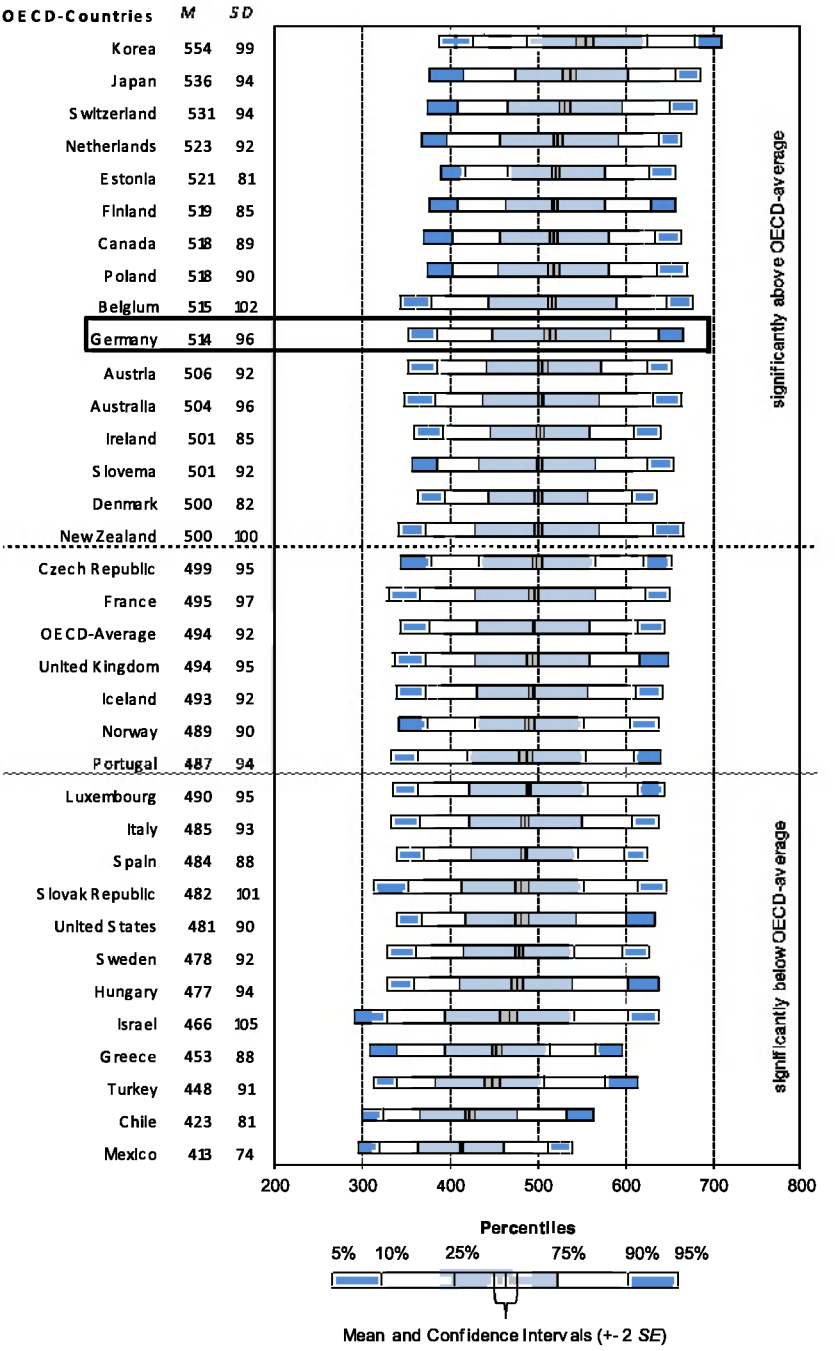
This article discusses the development of educational outcomes in PISA between 2000 and 2012. The question guiding the paper is: Which factors have contributed to a uniquely positive development and overall improvement of student performance? First, a few central results of PISA 2000 and PISA 2012 will be presented. Afterwards, three aspects of the deliberate change in Germany will be highlighted: a thorough diagnosis of the situation, an intense discourse between policy makers, the public and educational researchers as well as some specific examples of measures that have been taken. The article will be closed by an integrating discussion and implementations for the future.

PISA: Different Pictures of Educational Outcomes in Germany in PISA 2000 and PISA 2012

The picture revealed by PISA in Germany in 2001 yielded an overall achievement significantly below the OECD average in all three test domains. While the OECD average scores in all three domains were 500 in PISA 2000, Germany scored 484 in Reading, 490 in Mathematics and 487 in Science. Huge disparities in student performance were found according to social background, migration status and gender. Germany is by constitutional law divided into 16 federal states (so-called Länder), each holding its own responsibility for education. Surprisingly, an oversample of the PISA 2000 cohort showed enormous gaps between the highest and the lowest performing federal states (Baumert et al., 2002) within the country. This gap totalled 64 points on the PISA Mathematics scale, which corresponds to approximately two years of schooling (Ehmke et al., 2006). In other words, inside Germany, differences were found that covered nearly the range of OECD countries. Furthermore, the grading standards varied considerably between states, indicating that the feedback students received for their performance differed both according to the state they lived in and the school they attended within their state. For Germany, PISA 2000 provided data that had been unavailable before and hence allowed international benchmarking of characteristics of the educational system that shed light on the dimensions of statistical parameters that stood alone before PISA 2000. For example, the percentage of students who had repeated a school year in Germany was as high in no other country (Baumert et al., 2001; Krohne, Meier and Tillmann, 2004). Twenty four percent of all 15-year-old students had been retained and another 12 percent had started school one year later than usual, hence a total of 36 percent of students in Germany had a lagged school biography. PISA 2000 drew a picture of educational quality in Germany that was surprising and hence necessitated the need to get more data and reliable information about the situation.

Twelve years later, however, this picture looked considerably more positive (Prenzel, Sälzer, Klieme and Köller, 2013). Over the years, students in Germany have attained a level of performance that continuously exceeded the OECD average. Taking into account that the educational reforms and changes at the system level cannot be measured after just one or two PISA cycles, the development during more than a decade serves as a plausible indicator for monitoring progress. The positive trend in Germany is obvious, and at the same time the OECD average has not improved over the years. In part, this may be due to a changed combination of the

Figure 1: Percentile plots of mathematical competence in OECD countries



OECD countries; this group has been growing since PISA 2000, when it comprised 28 countries and PISA 2012, when it consisted of 34 countries. The constantly positive development of the average performance in Germany over five PISA rounds is quite unique among OECD countries, as the starting point (e.g., 490 points in Mathematics) was below the OECD average and still German students performed better and better in each of the following PISA rounds. The same applies to the other two domains, Reading (PISA 2000: M = 484; PISA 2012: M = 508) and Science (PISA 2000: M = 487; PISA2012: M = 524). Figure 1 shows the distribution of mean performance in Mathematics over all OECD countries in PISA 2012.

Besides an above average performance, the results of PISA 2012 showed that disparities due to social background and migration status have considerably decreased in Germany. The gain in student achievement goes back particularly to a large improvement in the achievement of low performing, low socioeconomic status and immigrant students. The number of students who do not surpass a level of proficiency that is worryingly low, has also decreased in size. Grade retention is still a common practice in Germany, but compared to 36 percent of students in PISA 2000 who had repeated a class at least once or started attending school later than usual, this percentage is now at 20 percent (Sälzer, Prenzel and Klieme, 2013; OECD average: 12.4 percent).

At this point, one can ask whether the improvement was obtained by merely focusing on student assessment and narrowing down the curriculum (Berliner, 2011), but also through enforced training or even drilling students (Ho, 2006). After all, PISA and other large-scale student assessments stand for a quite modern concept of educational 'efficiency' that can be measured through data collection (Carnoy, 2014). Taking into account non-cognitive outcomes also, PISA allowed analyses to control such assumptions. With regard to student characteristics and attitudes, we see that students in Germany have a high self-concept with regard to Mathematics (Schiepe-Tiska & Schmidtner, 2013) and they feel a strong sense of belonging towards their school (Sälzer et al., 2013). This had already been the case in PISA 2003, when Mathematics was the major domain of assessment for the first time. Problematic aspects such as late arrival or student absenteeism from school are, compared to the OECD average, marginal in German schools. As in the vast majority of participating countries, girls achieve significantly higher in reading competence than boys, while boys outperform girls in Mathematics (Gebhardt et al., 2013). While there are significant disparities between students with an immigrant background and those without one (Gebhardt et al., 2013), the relationship between

socio-economic background and student competence in Germany lies within the OECD average range (Müller and Ehmke, 2013). All in all, the results of PISA 2012 were positive news that revealed little cause for serious concern regarding a high price that is paid for improved performance, such as anxiety or lack of identification with school.

Looking back at the development between PISA 2000 and PISA 2012, the results of the latest PISA study seem somewhat impressive. Given that Germany had a difficult start with PISA, resulting in the well-known "PISA-shock", it is quite clear that good student performance is an achievement of a functioning educational system rather than a gift, a matter of the structure or curriculum. Numerous efforts have been taken, targeting the objective of improving student performance, reducing disparities and establishing a national concept of education rather than 16 Länder-specific educational systems. The following paragraphs highlight three central measures that we assume to have supported the positive development of Germany's performance in PISA.

A Through Diagnosis of the State of the Educational System

Along with the Third International Mathematics and Science Study (TIMSS), PISA 2000 was one of the first international large-scale student assessments that Germany participated in. For a long period of time, there was no doubt in Germany that the school system and its outcomes, especially in the domains of Mathematics and Science, were highly commendable. From the beginning, large scale assessment studies such as PISA have been conducted at research institutions in order to ensure high quality data analyses that go beyond a descriptive survey and provide both policy makers and the public with in-depth findings. These findings were meant to be used as a foundation for informed decisions or actions to be taken in order to improve educational quality in Germany.

The international comparisons as delivered by PISA provide benchmarks that help countries align themselves within a scale that indicates a country's position in an international context. How is a country's performance compared to other countries which have similar educational systems? How is it compared to countries with a fundamentally different structure? Without international benchmarks, many findings of studies within countries would have a less distinct meaning, for example 'high' satisfaction of teachers with their current job would stand-alone and one would be unable to tell whether this is good or bad news. In this regard, PISA provided large contributions to analysing the state of the educational system in Germany and to identifying its strengths and weaknesses.

From the beginning, numerous studies were attached to PISA in order to make sure the results of PISA would be useful for the German situation of a federal state that has 16 educational systems. National test instruments were administered as a supplement to the PISA tests in order to ensure an adequate measurement of nationally relevant curricular competences on the one hand and a valid interpretation of the international results of PISA on the other hand. Additional tests in Mathematics and Science were used to enhance the international design of PISA 2000, where both Mathematics and Science appeared as minor test domains for the first time. Reading was the major domain and was broadly investigated at the international level.

The German supplement was designed and administered so that already in the first cycle of PISA, a deep and differentiated analysis was possible as well as studying the association of international PISA items with items that were based on German curricula. These extra tests took place in a second testing day, along with additional tests capturing cross-curricular competences. Such overarching competences like communication, cooperation or school-related problem-solving were considered to be relevant indicators of students' readiness for life as a citizen and an autonomous part of society (Baumert et al., 2001).

Besides the cognitive part of PISA, researchers in Germany put an emphasis on student attitudes, experiences and beliefs as well. These constructs were measured using a student questionnaire, which in Germany was expanded by a number of additional scales and items. Among others, peer relationships were measured, and a nationally enhanced school questionnaire comprised of a number of questions that captured, for example, quality assurance and cooperation of schools with other institutions was used. A parent questionnaire had not been part of the international survey design of PISA 2000, but German researchers had developed one in order verify student responses with regard to their family background and in order to collect data on individual educational biographies.

In addition to the mentioned supplementary tests and questionnaires, the sample of PISA participants was extended as well. Taking into account the federal structure of Germany, politicians and researchers were interested in having a sample representative of the Länder. Such a sample would enable a thorough description of the situation within the Länder and comparing the results between the Länder. Germany actually conducted two PISA-studies: PISA-I (PISA International) and PISA-E (PISA Extended; Baumert et al., 2002; Prenzel, Baumert et al., 2005). While the international sample comprised between 200 and 250 schools, the sample representative at the Länder level consisted of around 1500 schools. Both

samples were overlapping, i.e. the PISA-I-schools were a subset of the PISA-E-schools. Besides this oversample at the school level, Germany enhanced the student sample within schools. Although the age of school enrolment in Germany corresponds to the OECD average (6 years; Sälzer et al., 2013), there are several measures commonly used in Germany which lead to quite a wide range of grade levels attended by 15-year-old students. Usually, the German PISA sample is drawn from students attending grade 7, 8, 9, 10 or 11. In order to get a more robust idea of the competences and characteristics of students towards the lower secondary level, an oversample of about ten students per 9th grade in each of the sampled schools was drawn. From PISA 2003 on, complete ninth grades (two per school) were drawn in addition to the age-based student sample of 15-year-olds (Prenzel et al., 2004). This additional sample was widely used to explore processes during lessons in order to analyse teaching and various interactions in classrooms (e. g., Seidel, 2002). One example for a comprehensive research and development undertaking is the so-called SINUS project, which will be described below.

At the same time, numerous research initiatives in the field of education have been activated. The German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) launched a Priority Programme focusing on educational quality in schools. The main focus of this programme was to explore school-related and outside school conditions for developing competences in Mathematics, Science and cross-curricular fields. One crucial aspect of the Priority Programmes is a nationwide cooperation between the participating researchers, particularly with regard to the federal structure of the country. Results from the Programme revealed possible reasons for the weaknesses of the German educational system that had recently been discovered by PISA and TIMSS and they investigated what could be done in order to improve educational outcomes in the future. More than 30 studies within the Programme took different perspectives on the educational quality in schools and focused on topics such as the effects of single-sex schools, evaluating training on self-regulated learning and problem-solving or instructional quality. All these topics were immediately related to aspects and issues that were identified to be either a weakness of the educational system in Germany or that were lacking reliable data and research findings.

To summarise, a lot of effort has been taken in Germany to improve insights into educational structures, processes and outcomes. The need for sound empirical research was clearer than ever after both TIMSS and PISA had discovered the devastating state of the educational system. Not only the participation in international comparisons, but also an extensive

national enhancement of these studies marked the beginning of a new era that did not settle for a common sense perception of educational effectiveness, but required reliable and solid evidence. Such evidence was the foundation for an informed discourse among policy makers, researchers and the public, as described in the following section.

Discourse Between Policy Makers, Researchers and the Public[1]

Collecting and analysing empirical data on educational institutions, processes and outcomes provided many institutional and political players with profound evidence that can help in different ways when decisions have to be made regarding the educational system. Doing this at an international level was quite new in Germany after PISA 2000. Policy makers, researchers and the public now had a common ground to start from, which could be used to prioritize the choice of options that were identified with regard to strengths and weaknesses of the German educational system at the secondary level. In this sense, PISA is said to have contributed to a rediscovery of internationally and nationally comparative education in the German discourse (Ertl, 2006). The public reaction to PISA was overwhelming in Germany. Hitting the headlines for weeks after the release of the results in December 2001, PISA entered everyday discussions and enduring debates on the quality of schools, teaching and teachers in Germany. Since then, education has made it into the focus of public attention where it used to be something that was taken for granted. Questions such as whether the school a child is attending matters for their achievement, development or well-being arose had rarely been asked before (Baumert, Trautwein and Artelt, 2003).

PISA affected Germany as a whole country. The problem revealed by PISA had to be addressed from a nationwide perspective. Germany had to face this problem in order to improve the quality of schools and educational outcomes. As education was, and still is, the responsibility of the Länder, this situation was new. And yet, Germany had to face this problem in order to avoid a repetition of the disastrous results in international LSA studies. The "PISA-shock" ignited a broad reform agenda in Germany, which is best seen in a document called *PISA 2000 – Central Fields of Action* by The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germa-

1 Referring to the 'public' in this article means the sphere outside formal politics, policy discourse or academia. In that regard, the 'public' comprises of citizens in a country, here: Germany, who discuss about educational issues from diverse perspectives, but not necessarily with a professional background or knowledge.

ny (KMK). This document is part of a long-term strategy (KMK, 2006) and states that Germany will participate in international LSA studies in the long run. Seven fields of action were identified that are to be transferred into specific measures to improve the educational situation in the whole country. Priority was given to (1) early support for improving language skills, (2) better links between pre-school and elementary school, (3) improvement of elementary schooling by a steady advancement of competences in reading, Mathematics and Science, (4) more support for children at risk, especially for those with an immigrant background, (5) measures of consequent advancement and quality assurance in teaching and school using nationwide standards and evaluation, (6) professional teacher education as well as (7) all-day schooling. One example of how these fields of action were addressed is the implementation of national educational standards. A group of researchers and experts in the field of education developed a framework for the development and implementation of standards (Klieme et al., 2003). This framework served as a kind of roadmap for combining societal objectives, scientific findings about competence development as well as concepts and procedures of test development.

Implementation of Overarching Measures

Besides a thorough diagnosis of problems within the educational system and a widespread discourse on education, the third consequence to be taken in Germany was to implement nationwide, overarching measures inside and across schools. One of these overarching and nationwide programmes, implemented after the TIMSS study in Germany, was the so-called SINUS-programme (*Enhancing the Efficiency of Teaching in Science and Mathematics*). SINUS was a model programme pursuing a sustainable improvement of the professional development of Mathematics and Science teachers. The programme suggested 11 modules as the core of SINUS which aimed to improve teaching and learning by advancing the development of a thorough change in Mathematics tasks (Prenzel, Friedrich & Stadler, 2009). The structure of modules enabled schools to select their own set of modules which could be combined and, later on, enhanced by new modules. This "new culture" of Mathematics tasks comprised of a much broader range of mathematical competencies (Niss, 2003) and focused on securing a basic understanding and fostering of cumulative learning in Mathematics. Teachers could find highly elaborated recommendations that helped them identify strengths and weaknesses in their own teaching as well as ideas and impulses for developing further approaches. SINUS intended to engage a large number of teachers who would commit to working in teams and use the modules to improve

tasks, materials and teaching approaches. It can therefore be considered as an example of teacher professional development (Oster Meier, Prenzel and Duit, 2010). Dissemination of approved successful modules took place first within the school, and then spread to other schools in the regional and later national school networks. The principle of SINUS was to encourage teachers teaching the same subject to cooperate within their school and, in the long run, between schools as a school network. Such cooperation would be coordinated at a local, a regional or a cross-regional level. In doing so, SINUS generated an enormous pool of materials available online to all interested teachers as well as in the form of manuals, books or teachers' magazines. Over the years, SINUS had developed from a starting size of 180 secondary schools in 1998 to 1750 schools in 2003, followed by a modified programme for primary schools.

SINUS was guided by five principles which best describe its approach. First, problematic areas should be handled by working on 11 so-called modules. These modules referred to an expertise which had identified certain problematic issues found in Mathematics and Science teaching in Germany (BLK Projektgruppe "Innovationen im Bildungswesen", 1997) and contained suggestions for developing one's teaching practices. Second, cooperation among teachers within the same school as well as within school networks (so-called school sets) was encouraged and fostered. Quality development and quality assurance was the third principle, meaning that schools were to establish routines in developing and maintaining quality-related factors. Fourth, the work of participating teachers was scientifically inspired and supported. Materials, professional development courses and counselling were meant to accompany the 11 modules mentioned. Finally, SINUS was enhanced by a number of research studies to delve into conceptual questions or questions of implementation, but also to evaluate the programme. Along with these five principles, SINUS pursued objectives at three different levels. Professional development of teachers was the first objective, meaning among others that teachers should learn to cooperate by sharing materials and test items, visiting each other during lessons or combining forces for preparing lessons. Cooperation has been found to be a crucial feature of effective schools (Sammons, 1999), but was not very common in German schools (Terhart, 2001) and should therefore be promoted. The second objective envisaged the level of teaching in lessons, which implied the need for improvement. Modules focusing on this second objective comprised of elements such as developing a new culture of tasks during lessons, scientific working or learning from mistakes (Prenzel, Carstensen, Senkbeil, Ostermeier and Seidel, 2005). The third objective is aimed at student learning

and achievement, seen as a learning outcome. Modules referring to this objective serve as a content-oriented framework for scaffolding in-depth learning, positive attitudes and high interest in Mathematics and Science.

SINUS schools were tested in 2003, using some of the PISA test items in order to compare student performance between SINUS schools and, as a control group, PISA schools (Prenzel, Carstensen et al., 2005). Results showed that students attending SINUS schools, at least in some school types, performed significantly better than students at non-SINUS (i.e. PISA) schools and also showed higher interest and more positive attitudes towards Mathematics and Science. To summarise, comprehensive schools and lower-secondary level schools (Hauptschule, offering a secondary-level I certificate after 9 years, HSA) profited the most from SINUS. This is in line with the findings of PISA 2012, where it turned out that the competence improvement over the past twelve years went back mostly to school types besides the Gymnasium, which are set up less academically than a Gymnasium. Therefore it seems as if especially those school-types with the most significant need for development have improved by participating in SINUS (Prenzel et al., 2004; Prenzel et al., 2006).

Another relevant overarching measure in Germany was national educational standards. Introducing national educational standards in Germany was one of the most fundamental measures to be taken as a consequence of the "PISA shock". In 2003 and 2004, the KMK published educational standards referring to primary level and secondary level I (e.g. KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik, 2003; KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik, 2004). At the primary level, the core of these standards referred to the school subjects of German language and Mathematics. At secondary level I, focus was set on German language, Mathematics, and the first foreign languages learned at school (English or, in some Länder, French). Standards were differentiated between a lower-secondary school-leaving certificate (HSA) after nine school years and an intermediate secondary school-leaving certificate (MSA), which is usually attained after ten school years (Pant et al., 2014).

For the first time, education was considered from a perspective of standards, competences and outcomes rather than from a curricular point of view and an input-oriented perspective. These standards being educational *objectives* means that the underlying concept of a young learner at a certain age is one that defines a competence that is expected from each individual who completes a certain level of schooling (Klieme et al., 2003). Such competences are described from a mastery point of view, i.e. they

indicate requirements towards teaching and learning at school and they name desirable learning outcomes for students. In that regard, Germany moved from an input-oriented perspective to an output-oriented point of view towards the educational system. Another important aspect that comes with this concept of educational standards is that they are verbalized and illustrated so teachers get a clear idea of what the curricular description of a standard or competence means. In addition, the description of standards can be understood as a framework for the development of tasks and tests (Klieme et al., 2003). Schools are expected to ensure that the competences defined in the national educational standards can be achieved by every student, regardless of social background. This idea shifts the role of schools from one that is responsible for successfully completing Länder-specific curricula to one that is responsible for ensuring the development of defined levels of competence that are valid across the Länder and across school-types. In order to sustainably develop and test the new educational standards in Germany, a specific research institution was founded in Berlin: the Institute for Educational Quality Improvement (IQB) at the Humboldt-Universität zu Berlin. One of the main responsibilities of the IQB is to administer sample-based comparative assessments of state-level (Länder) educational performance in order to assess the extent to which educational standards are being met across the 16 federal states of Germany. Along with several international Large-Scale Assessments such as PISA, TIMSS or PIRLS (Progress in International Reading Literacy Study), the so-called IQB National Assessment Studies examine student competences at the end of secondary level I within Germany. The first IQB National Assessment Study was conducted in 2009, in parallel with PISA 2009, and tested student competences in the subjects of German and the first foreign language at secondary level I (Köller, Knigge and Tesch, 2010). The next IQB National Assessment Study followed, along with PISA 2012, and examined student competences in Mathematics, Biology, Chemistry, and Physics (Pant et al., 2013).

One more aspect adds to the quite special situation in Germany's educational system. In Germany, there is only one secondary school-type that is prevalent throughout the country. All 16 Länder have the Gymnasium system, which qualifies students to attend tertiary education after completing eight or nine school years, respectively. So far, the PISA performance of Gymnasium students has been consistently high (Baumert et al., 2001; Klieme et al., 2010; Prenzel et al., 2004; Prenzel et al., 2007; Prenzel et al., 2013). As Gymnasium schools are the only school type that has not changed due to reforms over the PISA cycles and is prevalent in all federal states in Germany, only Gymnasium schools can be analysed from

a trend perspective. Therefore, it is not possible to describe the proficiency development of other school types over time. It can, however, be assumed that if the student performance at Gymnasium schools has not changed and the average performance of all students in Germany has significantly improved since PISA 2000, it must be due to an increase in student performance at other school types besides the Gymnasium (Prenzel et al., 2013). On one hand, this is good news, as the national educational standards published in 2003 and 2004 explicitly refer to the lower secondary and the intermediate secondary school-leaving certificate that can be obtained in school types other than Gymnasiums. So it seems that these school types have successfully implemented the objective lying behind the educational standards, namely to prepare students to transfer knowledge and processes to other, partly unknown situations and contexts and master a defined minimum of requirements that should be met by a typical young person leaving compulsory schooling. On the other hand, the fact that the highest achieving students who are grouped in Gymnasium schools have not improved their average performance from cohort to cohort as have students in other school types is not such good news. National educational standards with a focus on Gymnasium graduates (*Allgemeine Hochschulreife*, i.e. *Abitur* or high-school diploma) have recently been published and their effects will be visible in several years' time (e.g., KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik, 2012; KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik, 2013). Taking into account that the percentage of students per age cohort attending a Gymnasium has increased since PISA 2000 (from less than 30 percent to 36 percent) and still the average performance of this school type is consistently high, this expansion of the Gymnasium has been successful. However, the results of PISA 2012 also showed that high achieving students with good potential can and should be supported more in German schools, especially at the Gymnasium.

Both the SINUS programme and national educational standards tried to overcome boundaries between federal states in Germany and develop a spirit of cooperation and a common understanding of what education should pursue, how it could be done effectively and what good teaching actually means. Thinking of the teaching profession as one of the most responsible positions in society and with a lot of potential for professional development, SINUS chose a modular approach and encouraged teachers to overcome their being used to individualism and cooperate with their colleagues. At the same time, setting national educational standards reformed the understanding of proficiency, away from an

achievement-based orientation towards a mastery orientation. Mastery of educational standards, as understood in the German case, involves a regular assessment of student competences. These competences are defined in the educational standards rather than in the curriculum.

Summary and Discussion

This paper dealt with the recently published results of the fifth round of PISA in Germany and their meaning with regard to the development of the educational system since the "PISA-shock" that followed the first round of PISA in 2001. Comparing the pictures in 2001 and 2013, revealed that students in Germany have notably improved their average performance. It is at least in part for this reason that the positive results of PISA 2012 have been received quite happily in Germany. If it were not for the "PISA shock", results like those in PISA 2012, where students in Germany performed significantly above the OECD average in all domains, would not be very spectacular, but rather conforming to everybody's expectations. The paper then discussed how this quite uniquely positive development of student performance over five rounds of PISA could be explained. Three main aspects of the deliberate change in Germany have been elaborated: a thorough diagnosis of the state of the educational system, an intense discourse between policy makers, researchers and the public, as well as the implementation of overarching measures across the country.

With regard to the diagnosis PISA provided and the discourse it stimulated, Germany's educational system has benefited a lot from this study. Intense and, at times, heated debates have been important motivators the German development since PISA 2000. Knowing about problematic aspects within the system and struggling for better educational outcomes that are not achieved by a successful teaching of the test, but by merely by activating sustainable teaching, has proved to be healthy and fertile. However, PISA is limited with regard to what types of data and information it can deliver. PISA provides high-quality indicators of certain aspects of educational systems worldwide, such as specific strengths and weaknesses. It is thus a solid foundation for educational benchmarking at an international level. PISA does not however go beyond that. Solutions and strategies to implement changes have to be identified elsewhere. In PISA, countries can learn from one another if they wish, although some suggestions of publications in the PISA context miss the point by focusing too strictly on structural indicators of educational systems that may have some effects on student development and the explanation of variance, but do not necessarily take into account the underlying processes or country-specific coherences that go beyond explaining variance in student

performance (OECD, 2010). In Germany, some structural modifications have been undertaken, such as a reconsideration of the multi-tier secondary schooling system that is and has been prevalent in all federal states. However, these modifications have not touched the Gymnasium and the general structure of several secondary schooling types which children at age 10 are assigned to. For such reasons, flanking educational research is needed to enhance the potential that lies within PISA data. In Germany, this has been undertaken at the political level by the seven fields of action that have been stated by the Standing Conference of the Ministers of Education and at the academic level by accompanying research attached to PISA and by allocating the national project management to research institutions. If project management does not stop at the point where the collected data are submitted to another institution, but continues with country-specific analyses of PISA data, the value added to the participation in international large-scale assessments widely increases. Concerning the third pillar of deliberate change after PISA 2000, implementing overarching measures was the consequence that touched schools and teachers most directly. Introducing national educational standards initiated a shift from an input-oriented to an output- or outcome-oriented perspective in education. Measures to improve school quality, such as internal and external evaluation, standardised policies for Mathematics or teacher mentoring (OECD, 2013) are less prevalent in Germany than in most other OECD countries, but many schools already use these instruments and it can be assumed that more will follow when they hear about improvements and progress from their colleagues.

Can one conclude from this that Germany is done with PISA, now that student literacy is above the OECD average and many problematic aspects have been improved? Not quite yet. The uniquely positive development of Germany rests on several pillars: a nationwide strategy ('fields of action') bringing 16 educational systems together, national educational standards, overarching projects such as SINUS, but also a steady scientific evaluation and enhancement of large scale assessment studies and a continuous diagnosis of the state of the educational system. These aspects have become part of a regular public, scientific and political discourse. Such discourse was not prevalent before PISA in Germany and it now contributes to a common understanding of education as a goal to achieve and a project to work on. Education does not come by itself and it cannot be "provided" to students. Students find educational offers in schools and they have to make use of them in order to become educated young individuals, ready for a position in modern societies. Taking into account that the trend of PISA results (and, hence, of educational outcomes) in Germa-

ny is pointing upwards, there are also several challenges to be addressed. Although disparities have decreased, e.g. between immigrant and non-immigrant students or between students of high and low socio-economic background, these disparities still exist and need to be reduced further. Gender gaps, indicating that girls outperform boys in reading and boys outperform girls in Mathematics, have to be closed. This is especially relevant for Gymnasium schools, where many students achieve excellent results. It is within this group of top-achievers in Mathematics that girls are underrepresented . In general, strong and high achieving students require more support in Germany, e.g. through enrichment or specific school programmes offering a setting to develop excellent skills, but also high interest and motivation with regard to literacy skills as they are proposed and measured in PISA. With regard to supporting strong students as well as average or low achieving students, teaching practices and teaching quality need to be improved. Tasks that stimulate students to think, to be creative and interactive learners and to be able to transfer skills from the classroom to the real world are needed and these tasks have to be implemented in lessons. Given that other PISA participant countries have developed less positively than Germany, whether they have undergone structural reforms or not, this last aspect of teaching practices and teaching quality improvement seems to be more important by far than the question of comprehensive or tracked school systems providing a better learning environment for students. After all, it is not a matter of the system, but of the classroom processes when it comes to fostering student learning and preparing them for life as citizens in modern societies.

References

Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., and Weiss, M. (eds.) (2002) *Pisa 2000 - die Länder der Bundesrepublik Deutschland im Vergleich: [PISA-E].* Opladen: Leske + Budrich.

Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., and Weiss, M. (eds.). (eds.) (2001) *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich.* Opladen: Leske + Budrich.

Baumert, J., Trautwein, U. & Artelt, C. (2003) Schulumwelten -- institutionelle Bedingungen des Lehrens und Lernens. In: Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J., and Weiss, M. (eds.). *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland.* Opladen: Leske + Budrich, pp .261-331.

Berliner, D. (2011) Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*. 41(3), pp. 287-302.

BLK PROJEKTGRUPPE „INNOVATIONEN IM BILDUNGSWESEN" (1997) Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts". Materialien zur Bildungsplanung und zur Forschungsförderung: Vol. 60. Bonn: BLK, Geschäftsstelle.

Carnoy, M. (2014) Globalization, educational change, and the national state. In: Stromquist, N. P., and Monkman, K. (eds.) *Globalization and education. Integration and contestation across cultures*. Lanham, MD: Rowman & Littlefield., pp. 21-38.

Ehmke, T., Blum, W., Neubrand, M., Jordan, A. & Ulfig, F. (2006) Wie verändert sich die mathematische Kompetenz von der neunten zur zehnten Klassenstufe? In: Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. and Schiefele, U. (eds.) *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Müster: Waxmann, pp. 63-86.

Erth, 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), pp. 619-634.

Gebhard, M., Rauch, D., Mang, J., Sälzer, C. & Stanat, P. (2013) Mathematische Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In: Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (eds.) *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann, pp. 275-308.

Ho, E. S.-C. (2006) High stakes testing and its impact on students and schools in Hong Kong: What we have learned from the PISA studies. *Korean Journal of Educational Policy*. 3 (1), pp. 69-87.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E.,Vollmer, H. J. (2003) Zur Entwicklung nationaler Bildungsstandards: Expertise. Bonn: BMBF.

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (2003) Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (10. Jahrgangsstufe). Beschlüsse der Kultusministerkonferenz.

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (2004) Bildungsstandards im Fach Mathematik für den Hauptschulabschluss (9. Jahrgangsstufe). Beschluss vom 15.10.2004. Beschlüsse der Kultusministerkonferenz.

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (2006) *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring.* München: Wolters Kluwer.

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik (2012) Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.10.2012. Available from http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Deutsch-Abi.pdf (Accessed: 2nd May 2014).

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik. (2013) Konzeption zur Implementation der Bildungsstandards für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 10.10.2013. Available from: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/2013-10-10_Konzeption_Implementation_Bildungsstandards-AHR.pdf (Accessed: 2nd May 2014).

Köller, O., Knigge, M., and Tesch, B. (2010) *Sprachliche Kompetenzen im Ländervergleich.* Münster: Waxmann. Available from: http://www.e-cademic.de/data/ebooks/extracts/9783830923398.pdf (Accessed: 2nd May 2014).

Krohne, J. A., Meier, U., and Tillmann, K.-J. (2011) Sitzenbleiben, Geschlecht und Migration - Klassenwiederholungen im Spiegel der PISA-Daten. *Zeitschrift für Pädagogik.* 50 (3), pp. 373-391. Available from: http://www.pedocs.de/volltexte/2011/4816 (Accessed: 2nd May 2014).

Müller, K., and Ehmke, T. (2013) Soziale Herkunft als Bedingung der Kompetenzentwicklung. In: Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (eds.) *PISA 2012: Fortschritte und Herausforderungen in Deutschland.* Münster: Waxmann, pp. 245-274.

Niss, M. (2003) Mathematical competencies and the learning of Mathematics: The danish KOM project. In: Gagatsis, A., and Papastavridis, S. (eds.) *3rd Mediterranean conference on mathematical education.* Athens, Greece: Hellenic Mathematical Society and Cyprus Mathematical Society, pp.115-124.

OECD (2010) *What Makes a School Successful? PISA 2009 Results.* Paris: OECD Publishing.

OECD (2013) *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices. Volume IV.* Paris: OECD Publishing.

Ostermeier, C., Prenzel, M., and Duit, R. (2010) Improving Science and Mathematics instruction - The SINUS-Project as an example for re-

form as teacher professional development. *International Journal of Science Education*. 32 (3), pp. 303-327.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., and Pöhlmann, C. (eds.) (2013) *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., and Pöhlmann, C. (2014) *The IQB National Assessment Study 2012: Summary*. Münster: Waxmann.

Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., and Pekrun, R. (eds.) (2007) *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. and Schiefele, U. (eds.) (2004) *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs*. Münster, New York: Waxmann.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. and Schiefele, U. (eds.) (2005) *PISA 2003. Der zweite Vergleich der Länder in Deutschland -- Was wissen und können Jugendliche?* Münster: Waxmann.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. and Schiefele, U. (eds.) (2006) *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Müster: Waxmann.

Prenzel, M., Carstensen, C. H., Senkbeil, M., Ostermeier, C., and Seidel, T. (2005) Wie schneiden SINUS-Schulen bei PISA ab? *Zeitschrift für Erziehungswissenschaft*. 8 (4), pp. 540-561.

Prenzel, M., Friedrich, A., and Stadler, M. (eds.) (2009) *Sinus-Transfer. Von SINUS lernen: Wie Unterrichtsentwicklung gelingt (1st ed.)*. Seelze: Kallmeyer.

Prenzel, M., Sälzer, C., Klieme, E., and Köller, O. (eds.) (2013) *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.

Sälzer, C., Prenzel, M., and Klieme, E. (2013) Schulische Rahmenbedingungen der Kompetenzentwicklung. In: Prenzel, M., Sälzer, C., Klieme, E., and Köller, O. (eds.) *PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann, pp.155-187.

Sammons, P. (1999) *School Effectiveness: Coming of age in the twenty-first century*. Lisse: Swets & Zeitlinger.

Schiepe-Tiska, A., and Schmidtner, S. (2013) Mathematikbezogene emotionale und motivationale Orientierungen, Einstellungen und Verhaltensweisen von Jugendlichen in PISA 2012. In: Prenzel, M., Sälzer, C., Klieme, E., and Köller, O. (eds.) *PISA 2012: Fortschritte und Herausforderungen in Deutschland.* Münster: Waxmann, pp.100-121.

Seidel, T. (2002) *Lehr-Lernskripts im Unterricht. Freiräume und Einschränkungen für kognitive und motivationale Prozesse beim Lernen -- Eine Videostudie im Physikunterricht.* Münster: Waxmann.

Terhart, E. (2001) *Lehrerberuf und Lehrerbildung: Forschungsbefunde, Problemanalysen, Reformkonzepte.* Weinheim: Beltz.

The Influence of PISA on Educational Policy in Canada: Take a Deep Breath

Pierre Brochu

The Programme for International Student Assessment is the most widely recognized international assessment of its kind. Designed to test student knowledge and skills in the core subject areas of reading, mathematics, and science, PISA provides policy-oriented indicators of the basic competencies of youth before the end of compulsory schooling.

In just over 10 years, PISA has become the ultimate reference for international, large-scale assessment. Its influence in government, in the education community, in the media, and with the general public has grown exponentially since 2000. In addition, by 2015, each of the major domains will have been tested twice, thus enabling participating countries not only to compare themselves with other education systems but also to compare their own results over time, thanks to the nine-year cycle (Table 1).

Table 1: PISA Assessment Cycle

	2000	2003	2006	2009	2012	2015[1]	2018[2]
Major Domain	Reading	Mathematics	Science	Reading	Mathematics	Science	Reading
Minor Domains	Mathematics Science	Reading Science	Reading Mathematics	Mathematics Science	Reading Science	Reading Mathematics	Mathematics Science
Supplemental Domain		Problem Solving	Computer-Based Science	Computer-Based Reading	Computer-Based Problem Solving	Collaborative Problem Solving	To be determined

1 See http://www.oecd.org/pisa/pisaproducts/pisa2015draftframeworks.htm
2 See http://www.oecd.org/pisa/pisaproducts/PISA-2018-documents-for-bidders.htm

The objective of this paper is to demonstrate how PISA results can be used in Canada — where multiple ministries/departments are responsible for education — by drawing on the experiences of selected countries, most of which have a similar federal structure. It presents cases studies of how several countries have reacted and responded to PISA results over the last decade to plan education reform and improve education. Firstly, I will briefly describe the particularity of the PISA administration in Canada and summarize the Canadian results since 2000. Secondly, I will provide several examples of how other countries with either a similar political structure or similar challenges have used their PISA results to implement the changes necessary to respond to these results. Thirdly, I will make some observations on the growing use of trend data related to PISA and how Canadian provinces have responded to the PISA results thus far. Finally, I will argue that, among the lessons learned from PISA, Canadian education systems can derive the greatest benefit from the experiences in their own country as well as from those in other countries by considering a long-term perspective in order to orient educational policy in a constructive and responsible manner.

PISA in Canada

Canada has participated in PISA since its inception in 2000. Approximately 20,000 students from 1,000 schools in the 10 provinces are typically evaluated (Brochu et al., 2013). This sample size is significantly larger than that of most other countries, owing to the jurisdictional and linguistic make-up of Canada's education systems. Education is the exclusive responsibility of the provinces and territories in Canada and is delivered in both English and French. A large sample is thus required in order to provide statistically reliable results for every education system in the country, as well as for both language groups.

Results of the initial PISA survey in 2000 were, for many countries, surprising. As will be explained later, some countries, such as Finland, South Korea, Japan, and Canada, performed well above the OECD average, which was duly noted by the media. Others, however, such as Germany, the United Kingdom, and the United States, fell far short of where they expected to be in the world's education league tables.

In Canada, the overall results of PISA have generally been reassuring, as its average scores have regularly been well above the OECD averages (Table 2); in fact, only Finnish students have consistently outperformed Canadian 15-year-olds since 2000, although Asian economies have taken over the lead most recently (OECD, 2013a).

Table 2: PISA 2000 to 2012 Average Score (standard error), Canada and the OECD

	2000		2003		2006		2009		2012	
	OECD	Canada	OECD	Canada	OECD	Canada	OECD	Canada	OECD	Canada
Reading	500 (0.6)	534 (1.6)	494 (0.6)	528 (1.6)	492 (0.6)	527 (1.6)	493 (0.5)	524 (1.2)	496 (0.5)	523 (1.2)
Mathematics	...	...	500 (0.6)	532 (1.8)	498 (0.5)	527 (2.1)	496 (0.5)	527 (2.6)	494 (0.5)	518 (2.7)
Science	...	...	...	...	500 (0.5)	534 (2.0)	501 (0.5)	529 (3.0)	501 (0.5)	525 (1.9)

Canada also stands out not only for attaining high results but also because of the considerable equity in achievement (OECD, 2011b), as shown in Table 3. The country has been cited as a model for permitting students to reach their full potential as constructive and reflective citizens regardless of the school they attend (OECD, 2013b), as demonstrated by the many measures of equity used by PISA: a low proportion of low achievers; a relatively small achievement gap between high and low achievers; a small proportion of variance explained by between-school differences; a weak relationship between performance and socioeconomic status; and a small gap between students from an immigrant background and those born in the country (Levin, 2012).

Table 3: PISA 2012 Mathematics Selected Measures of Equity, Canada and the OECD

	Canada	OECD
Proportion of 15-year-olds below level 2	14%	23%
Gap between 90th and 10th percentiles	231 points	239 points
Proportion of total variation explained by between-school variance	18.4%	36.8%
Percentage of variance explained by socioeconomic status	9.4%	14.6%
Gap between non-immigrant and first-generation immigrant students	-6 points	−45 points

These results are particularly interesting given that Canada is the only OECD country without a central (federal) ministry/department of education; since, by definition, centralization can facilitate the creation of equitable education policies and help to ensure equitable resource allocation. Other countries with a federal presence in education such as the United States or Germany (discussed below) have generally achieved average performance on PISA since 2000 with far less equity than Canada. However, as argued by Wallner (2013), the high degree of equity in Canada may well be a *consequence* of decentralization, as the Canadian sys-

tems allow provinces and territories to adapt their policies, curricula, and resource allocation to the specific needs of their populations. That being said, there is a measure of equity for federal countries that has not drawn a lot of attention internationally, but does warrant a closer look for Canada. In 2000, the gap in reading between the lowest and highest achieving provinces was 49 points. In 2012, the gap was 45 points, suggesting slightly more equity. However, this equity came at the cost of achievement, in that the highest achieving provinces reached 15 points less than in 2000 and the lowest achieving province, 6 points less.

PISA Results in Other Countries

Other federal countries participating in PISA have generally shown results much lower than Canada's. Germany, the United Kingdom, the United States, and Spain have all seen their results close to the OECD average, while Australia and New Zealand have been slightly above.

Since the mid-1950s, Germany has stood out as a world leader in higher education and as one of a handful of countries where compulsory education has been well established for the past half century (UNESCO, 2000). However, the initial PISA 2000 results created what has since been referred to as "PISA shock." The OECD PISA 2000 ranking had a huge impact in the country to the point where it "stopped the complacency and self-confidence with which Germany had looked at its education system for too long" (*Der Spiegel*, as cited in Dräger, 2012, p. 5). Facing results that placed the country below the OECD average, both orders of government (federal and *länder*) proposed urgent reforms, which focused on outputs and Germany's international competitiveness (Martens and Niemann, 2010) and laid great emphasis on empirical research and pedagogic practice (Ertl, 2006). They included a significant increase in student testing, changes to curricula, increases in funding, and additional measures of quality control (Grek, 2009; Anderson, Chiu and Yore, 2010; Neumann, Fisher and Kauertz, 2010). Interestingly, the most recent PISA results (2012) have confirmed the significant improvement in Germany's PISA average scores and more equity in education outcomes (OECD, 2013c). It is worth noting, however, that the streaming of students (a notable feature of the German education system that has been strongly criticized in some quarters) remains untouched.

The earlier PISA results also triggered strong reactions in the United Kingdom. While the UK's participating entities (England, Scotland, and Northern Ireland) each registered areas of positive outcomes, the results were less encouraging in aggregate. Since the first round of PISA, the UK's performance has been portrayed as "at best stagnant, at worst de-

clining" (Chakrabarti, 2013; Coughlan, 2013), with teacher qualifications and school autonomy being given, among others, as possible reasons for the lower achievement; nonetheless, this did not lead to concrete policy change (Baird et al., 2011). However, it has also been argued that PISA was a catalyst for an increase in testing with an explicit reference to PISA-related performance targets in Ireland (Breakspear, 2012; Figazzolo, 2009).

In the United States, however, the "very average" results from the early rounds of PISA were largely ignored by the American education community, policy makers, and the media. This may have been due, in part, to the fact that the PISA sample was relatively small: like Canada, the United States maintains a decentralized education system (albeit with a significant federal presence), and the PISA samples did not yield results that could be analysed at a state level.

While little has actually been done to reform education in America based on PISA findings, more attention is being paid to them, as seen by the positions taken by the U.S. Secretary of Education (Duncan, 2013) and expressed at the recent International Summit on the Teaching Profession. PISA-related discourse in the U.S. has been not on "spending more" but on "spending more wisely," in recognition of the fact that the United States is second only to Luxembourg in terms of per-student spending on education (Paine and Schleicher, 2011). Specifically, in an extensive comparative analysis of PISA results in the United States with those in high performing countries, it has been argued that resources need to be redirected to socioeconomically disadvantaged schools (Merry, 2013), teacher salaries (OECD, 2011a), and programs that increase teacher effectiveness (Hanushek, in Froese-Germain, 2010, p. 18).

In Spain, results have been characterized by lower achievement and lower equity both between regions and between sub-populations with no tangible improvement over time (OECD 2013d).

Identifying the factors that drive PISA results in high-performing countries is difficult (OECD, 2011a). Education systems are highly complex, and virtually any combination of their elements can be cited in explaining PISA results (whether strong or weak). As explained by Figazzolo (2009), "Taken as they are, which is (...) very often what happens, PISA results can be used to support A as well as the opposite of A" (p. 28). Thus, it is advisable not to look at systems or factors in isolation, but rather to consider how a combination of factors works to produce high performance — and whether this combination of factors can be replicated in other similar contexts. This has clearly been how the OECD has elected to portray individual country results and how many countries have used PISA results to further their political agendas. The pressure created by PISA about learn-

ing from the best (i.e., highest-performing countries based on PISA) has triggered the emergence of a new phenomenon labelled "educational tourism" (Robelen, 2013), where high-performing countries are visited by delegations from lower-performing countries.

A case in point is Finland. Its leading performance in PISA since the initial round has generated an exceptional amount of interest and has been attributed to a variety of factors: non-differentiation (i.e., no tracking or streaming of students); highly qualified and respected teachers; the absence of high-stakes national assessments; and a decision-making process for curriculum and teaching approaches that is decentralized and school-based (Valijarvi et al., 2002; Malaty, 2012). Obviously, other countries have these factors in place to some extent, but, as with any good recipe, the Finnish secret lies in having the right ingredients, in the right amount, and in the right context.

Not surprisingly, many education stakeholders from around the world wished to emulate Finland's results after the first round of PISA. However, they tended to focus on those factors that furthered specific political, educational, or economic agendas. Teacher unions, for example, cited the absence of a testing regime or the presence of highly educated teachers in Finland (OSSTF, 2007; Figazzolo, 2009). Other stakeholders pointed to different factors, such as the absence of streaming (as compared to Germany, among others) (Ammermüller, 2005); a homogeneous population (Entorf and Minoiu, 2005); the flexible curriculum and school structure (OECD, 2011a); and the late entry point for compulsory schooling (Mead, 2008).

The Finnish model contrasts with another successful system, namely that of South Korea, which has been used to justify very different policies (Pearson, 2012). These include long study hours (Chakrabarti, 2013); private investment in education (Lloyd, 2012); a combination of high expectations and a curriculum that emphasizes creativity and problem solving (Marginson, 2013); and (unlike Finland) a strong culture of testing (Dalporto 2013).

There are also countries where results have been fluctuating over time. This is the case in Japan, whose stellar results in the initial round of PISA were followed by a decrease in reading in 2003 and another decrease in mathematics in 2006. Japanese policy makers responded swiftly to these declines by initiating a multi-year plan to improve reading followed by the implementation of national tests and a national curriculum review (Ninomiya and Urabe, 2011).

Trends: The New Yardstick in Education Effectiveness

During the first few cycles of PISA, country ranking was the most commonly reported result in the media (as opposed to average score or the proportion of students at certain performance levels). In 2009, however, the emphasis started to shift, with greater interest being given to changes in a country's results over time. This new focus was in part a response to the changes in the countries participating in PISA from one cycle to the next: with new (often high-performing) countries and economies joining in later cycles, it became increasingly difficult for established countries to make sense of their "ranking" over time.

To assist in better evaluating changes within a country, the OECD developed an index of annualized change in performance that takes into account the number of years between each measure. This was provided in the 2012 International Report (OECD, 2013a) and gave participating countries a robust indicator of internal progress (or decline) over time.

According to this index (see Table 4), about half of all countries have improved over time in reading since 2000; a number of countries have seen a decrease in mathematics performance since 2003; and a majority of countries have remained stable in science since 2006.

Table 4: Number of countries and economies and change in average score over time (PISA 2000-2012)

Reference Year (2012)		Reading 2012 compared to 2000	Mathematics 2012 compared to 2003	Science 2012 compared to 2006
	–	32	25	19
		22	25	37
	+	19	14	8

Note: (author's calculations)
+ Number of countries and economies where average score has increased over time
Number of countries and economies where average score has not changed significantly over time
– Number of countries and economies where average score has decreased over time

As a result of the country's high performance since 2000, reactions to PISA results have been more muted in Canada than in many other countries. This has extended to the policy area, where policy makers have steered clear of making drastic changes based on limited data or research (Hess, 2008). However, provinces that did not fare as well as expected have introduced some moderate initiatives, based in part on PISA data and which could be characterized as fine-tuning an already strong system. New Brunswick reconsidered its French-immersion program (Dicks, 2008; Cooke, 2009) and Ontario launched both its *Literacy and Numer-*

acy Initiative (to improve reading and mathematics results at the primary level) and its *Student Success Initiative* (to increase the high school graduation rate) (OECD, 2011a).

A recent decline in national results, however, may stimulate a stronger response. Described in some quarters as "a national emergency" (*Globe and Mail*, 2013), the weakening of mathematics skills evident in PISA 2012 has created calls for immediate action in two areas: the training of teachers in mathematics and a review of the content of provincial mathematics curricula.

Teacher training was, interestingly enough, often cited as a reason for Canada's strong results in early PISA cycles (OECD, 2004). More recently, however, the international Teacher Education and Development Study in Mathematics (TEDS-M), administered in 2008 (CMEC, 2010), pointed out that many Canadian elementary-level teachers lacked knowledge in mathematics and mathematics pedagogy, while those at the lower-secondary level lacked training in assessment. The study also noted that a smaller proportion of mathematics educators (those teaching future teachers) in Canadian universities were specialized in mathematics at the doctoral level compared to the international average.

Provincial mathematics curricula came under scrutiny from a number of observers for their emphasis on "new math." This approach lies at the heart of mathematics curricula in most of Canada (with Quebec, whose mathematics scores greatly outranked the rest of the country, a notable exception) and was singled out for favouring discovery learning and problem solving over "basic knowledge and skills" and "daily-use math" (Alphonso, 2013).

Lessons Learned

At the time of writing this article, several provinces are considering proposals in the two areas of teacher training and curriculum renewal (Alberta Education, 2014; British Columbia Education, 2013; Manitoba Education, 2013; Nova Scotia Education and Early Childhood Development, 2013; Ontario Ministry of Education, 2014), although many initiatives had already been undertaken prior to the release of the PISA 2012 results.

Canadian provinces would be well advised to reflect carefully on any reforms they may undertake. Judging by the situation in other countries, there appears to be an inclination to push the panic button and implement reforms based on limited evidence. Too often, causation and correlation are confused when discussing PISA results (Mortimore, 2009), and any outcomes should be validated with other data sources, such as

other international studies and pan-Canadian or provincial results. Furthermore, the growing use of trend data, rather than reliance on comparative rankings alone, can significantly improve the usefulness of PISA results, in particular in those countries such as Canada, the United States, Germany, Italy, or Spain, where PISA results are available at the regional, state, or provincial/territorial level.

The ultimate goal of education should not be to finish first in the PISA race or improve in international rankings. Instead, education should enhance performance levels and equity for *all* students (Yore et al., 2010). PISA indicators should be used to attain these goals, by being integrated into national/federal policies and practices (Breakspear, 2012), not by replacing them.

Reform of Canada's education systems should acknowledge that from an international perspective, the country is still regarded as a model to emulate. It would make little sense to implement major changes to education policies across Canada based solely on PISA results when other countries use the same results to justify emulating Canada.

Canada also benefits from a very large sample size in PISA, which allows results to be analysed at a fine-grained level. Canada's provinces can thus learn not only from the experience of other countries, but also from their neighbouring provinces (Wallner, 2013). In the case of the most recent mathematics results, for example, Quebec appears to have much to impart to the rest of the country, as its results place it among the highest of all PISA participants in the world. Not only is it one of those jurisdictions that elected not to completely redesign its mathematics curriculum to integrate discovery learning into the program of study (Alphonso, 2013), it is also the province with the most mathematics teachers who are actually specialized in teaching mathematics (CMEC, 2012) and the only province where the proportion of low achievers in mathematics has not increased over the past nine years (Brochu et al., 2013). As another example, students in British Columbia have achieved sustained high performance in reading, science, and problem solving in recent PISA cycles, and that province's on-going curriculum review is cited as one of the reasons for their success (Sherlock, 2014).

This paper has attempted to analyse the impact of PISA in several countries where PISA results have garnered considerable attention over the past decade. It argues that silver bullets based on PISA results are not only unrealistic but should be avoided (Hargreaves & Shirley, 2012). Education systems are complex entities requiring a thoughtful, systematic, and balanced approach to reform (Sahlberg, 2011).

References

Alberta Education (2014) *Curriculum redesign - Letter from the Minister.* Available from: http://education.alberta.ca/department/ipr/curriculum/ministercrletter.aspx (Accessed 14th March 2014).

Alphonso, C. (2013, 3rd December) Canada's falling math rankings renews push for national standards. *Globe and Mail.* Available from: http://www.theglobeandmail.com/news/national/education/canadas-falling-math-rankings-renews-push-for-national-standards/article15755434/ (Accessed 3rd December 2013).

American Psychological Association (2010) *Today's superheroes send wrong image to boys, say researchers* [Press release]. Available from: http://www.apa.org/news/press/releases/2010/08/macho-stereotype-unhealthy.aspx (Accessed 11th December 2014).

Ammermüller, A. (2005) PISA: What makes the difference? Explaining the gap in PISA test scores between Finland and Germany. *Discussion Paper No. 04-04.* Centre for European Economic Research. Available from: ftp://ftp.zew.de/pub/zew-docs/dp/dp0404.pdf (Accessed 11th December 2014).

Anderson, J. O., Chiu, M.-H., and Yore, L.D. (2010) First cycle of PISA (2000-2006) – International perspectives on successes and challenges: Research and policy directions. *International Journal of Science and Mathematics Education. 8,* pp. 373-388.

Anonymous (2007, December 21) PISA 2006 - The politics of testing. *Update.* 35 (4). Toronto: Ontario Secondary School Teachers' Federation.

Baird, J.-A., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., and Daugherty, R. (2011) *Policy effects of PISA.* Oxford University Centre for Educational Assessment. Available from: http://oucea.education.ox.ac.uk/wordpress/wp-content/uploads/2011/10/Policy-Effects-of-PISA-OUCEA.pdf (Accessed 11th December 2014).

Breakspear, S. (2012) The policy impact of PISA. An exploration of the normative effects of international benchmarking in school system performance. *OECD Education Working Papers, No. 71.* Paris: OECD Publishing.

British Columbia Education (2013) *B.C. students among best internationally.* [Press release]. Available from: http://www.newsroom.gov.bc.ca/2013/12/bc-students-among-best-internationally.html (Accessed 11th December 2014).

Brochu, P., Deussing, M-A., Houme, K., and Chuy, M. (2013) *Measuring up: Canadian results of the OECD PISA Study. The performance of*

Canada's youth in mathematics, reading and science. 2012 first results for Canadians aged 15. Council of Ministers of Education, Canada.

Chakrabarti, R. (2013, 2nd December) South Korea's schools: Long days, high results. *BBC News.* Available from: http://www.bbc.co.uk/news/education-25187993 (Accessed 11th December 2014).

Cooke, M. (2009) *PISA 2009: It's not about the Leaning Tower.* New Brunswick Teachers' Association. Available from: http://www.nbta.ca/profession/curriculum/finance14.html (Accessed 11th December 2014).

Council of Ministers of Education, Canada (2010) *Teacher education and development study in mathematics 2008: Canadian report.* Toronto: Author.

Council of Ministers of Education, Canada (2012) *Contextual report on student achievement in mathematics.* Toronto: Author.

Coughlan, S. (2013, December 3) Pisa tests: UK stagnates as Shanghai tops league table. *BBC News.* Available from: http://www.bbc.com/news/ (Accessed 11th December 2014).

Dalporto, D. (2013) South Korea's school success. *We are teachers.* Available from: http://www.weareteachers.com/hot-topics/special-reports/teaching-around-the-world/south-koreas-school-success (Accessed 11th December 2014).

Dicks, J. (2008) *The case for early French immersion: A response to J. Douglas Willms.* Second Language Research Institute of Canada: University of New Brunswick. Available from: www.unbf.ca/l2 (Accessed 11th December 2014).

Dräger, J. (2013) *Accountability as a driver for reform: The "PISA shock" of 2001 – A spotlight on the case of Germany.* Harvard University, July 26th, 2012. Available from: http://www.hks.harvard.edu/pepg/conferences/July_2012_Presentations/Drager_Panel%201.pptx (Accessed 11th December 2014).

Duncan, A. (2013) *The threat of educational stagnation and complacency. Remarks of U.S. Secretary of Education Arne Duncan at the release of the 2012 Programme for International Student Assessment (PISA)* Available from: http://www.ed.gov/news/speeches/threat-educational-stagnation-and-complacency (Accessed 11th December 2014).

Entorf, H., and Minoiu, N. (2005) What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration. *German Economic Review.* 6(3), pp.355-376.

Erth, 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), pp. 619-634.

Figazzolo, L. (2009) *Impact of PISA 2006 on the education policy debate.* Education International. Available from: http://download.ei-ie.org/docs/IRISDocuments/Research%20Website%20Documents/2009-00036-01-E.pdf (Accessed 11th December 2014).

Froese-Germain, B. (2010) *The OECD, PISA and the impacts on educational policy.* Canadian Teachers' Federation. Available from: http://www.ctf-fce.ca (Accessed 11th December 2014).

Grek, S. (2009) Governing by numbers: the PISA 'effect' in Europe. *Journal of Education Policy.* 24 (1), pp. 23-37.

Hargreaves, A., and Shirley, D. (2012) The international quest for educational excellence: Understanding Canada's high performance. *Education Canada*, Fall 2012, pp. 10-13.

Hess, F.M. (2008) The politics of knowledge. *The Phi Delta Kappan.* 98 (5), pp. 354-356.

Levin, B. (2012) Greater equity in education. *Phi Delta Kappan.* 93 (7), pp.74-76.

Lloyd, M. (2012, May 14) What we can learn from South Korea? *Edutech Associates.* Available from: http://edutechassociates.net/2012/05/14/what-can-we-learn-from-south-korea (Accessed 11th December 2014).

Malaty, G. (2012) *PISA results and school mathematics in Finland: strengths, weaknesses and future.* Available from: http://www.aka.fi/fi/A/Suomen-Akatemia/Blogit/Markku-Leskela-PISA-ja-olympialaiset/Asia-on-tutkittu-httpmathunipait~grim21_project21_charlotte_MalatyPaperEditpdf/ (Accessed 11th December 2014).

Manitoba Education (2013, 3rd December) *Province's smaller classes initiative results in 130 new teachers in schools this year.* [Press release] Available from: http://news.gov.mb.ca/news/index.html?archive=&item=19797 (Accessed 11th December 2014).

Marginson, S. (2013, 16th December) Australia falls further behind in PISA test of basic education skills. *Higher Education.* Available from: http://www.theaustralian.com.au/higher-education/australia-falls-further-behind-in-pisa-test-of-basic-education-skills/story-e6frgcjx-1226783857196 (Accessed 11th December 2014).

Martens, K., and Niemann, D. (2010) *Governance by comparison — How ratings & rankings impact national policy-making in education. TranState Working Papers, No. 39.* Available from: http://www.staatlichkeit.uni-bremen.de (Accessed 11th December 2014).

Mead, S. (2008) How Finland educates the youngest children. *The Early Ed Watch Blog.* Available from: http://www.newamerica.net/blog/early-ed-watch/2008/how-finland-educates-youngest-children-9029 (Accessed 11th December 2014).

Merry, J.J. (2013) Tracing the U.S. deficit in PISA reading skills to early childhood: Evidence from the United States and Canada. *Sociology of Education*. 86 (3), pp. 234-252.

Mortimore, P. (2009) *Alternative models for analysing and representing countries' performance in PISA*. Education International Research Institute. Brussels.

Ninomiya, A., and Urabe, M. (2011) Impact of PISA on education policy - The case of Japan. *Pacific-Asian Education*. 23 (1), pp. 23-30.

Nova Scotia Education and Early Childhood Development (2013) *Nova Scotia students perform well in international assessments* [Press release]. Available from: http://novascotia.ca/news/release/?id=20131203001 (Accessed 11th December 2014).

Neumann, K., Fisher, H.E., and Kauertz, A. (2010) From PISA to educational standards: The impact of large-scale assessments on science education. *International Journal of Science and Mathematics Education*. 8, pp. 545-563.

OECD (2004) *What makes school systems perform? Seeing school systems through the prism of PISA*. Paris: OECD Publishing.

OECD (2010) *PISA 2009 results: What students know and can do. Student performance in reading, mathematics and science (Volume I)*. Paris: OECD Publishing.

OECD (2011a) *Strong performers and successful reformers in education lessons from PISA for the United States*. Paris: OECD Publishing.

OECD (2011b) Education: Bridging the classroom divide. *OECD Observer*. No. 284.

OECD (2013a) *PISA 2012 results: What students know and can do (Volume I): Student performance in mathematics, reading and science*. Paris: OECD Publishing.

OECD (2013b) *PISA 2012 results: Excellence through equity: Giving every student the chance to succeed (Volume II)*. Paris: OECD Publishing.

OECD (2013c) *Programme for International Student Assessment (PISA) results from PISA 2012. Country note: Germany*. Available from: http://www.oecd.org/pisa/keyfindings/PISA-2012-results-germany.pdf (Accessed 11th December 2014).

OECD (2013d) *Programme for International Student Assessment (PISA) results from PISA 2012. Country note: Spain*. Available from: http://www.oecd.org/pisa/keyfindings/PISA-2012-results-spain.pdf (Accessed 11th December 2014).

Ontario Ministry of Education (2014) *New math supports and resources for the classroom Ontario government committed to student success* [Press release]. Available from: http://news.ontario.ca/edu/

en/2014/01/new-math-supports-and-resources-for-the-classroom.html (Accessed 11th December 2014).

Paine, S.L., and Schleicher, A. (2011) *What the U.S. can learn from the world's most successful education reform efforts. Policy Paper: Lessons from PISA.* McGraw-Hill Research Foundation. Available from: http://hub.mspnet.org/index.cfm/22436 (Accessed 11th December 2014).

Pearson (2012) *The Learning Curve: Lessons in country performance in education. 2012 Report.* Available from: www.thelearningcurve.pearson.com (Accessed 11th December 2014).

Robelen, E. (2013, 18th December) Finland's global standing in education takes hit with latest PISA results. *Education Week.* Available from: http://blogs.edweek.org/edweek/curriculum/2013/12/educational_tourism_to_finland.html (Accessed 11th December 2014).

Sahlberg, P. (2011) PISA in Finland: An education miracle or an obstacle to change? *CEPS Journal.* 1 (3), pp.119-140.

Sherlock, T. (2014, April 1). B.C. students among best in world at problem solving but 'work' remains. *The Vancouver Sun.* Available from: http://www.vancouversun.com/technology/students+among+best+world+problem+solving+work+remains/9692940/story.html (Accessed 11th December 2014).

UNESCO (2000) The world education report 2000. *The right to education: towards education for all throughout life.* UNESCO, Available from: http://www.unesco.org/education/information/wer/ (Accessed 11th December 2014).

Valijarvi, J., Linnakyla, P., Kupari, P., Reinikainen, P., and Arffman, I. (2002) *The Finnish success in PISA - and some reasons behind it: PISA 2000.* Finnish Institute for Educational Research. University of Jyvaskyla.

Wallner, J. (2013, 6th December) Math scores are down. Don't panic. *Ottawa Citizen.* Available from: http://www.ottawacitizen.com/opinion/op-ed/Math+scores+down+panic/9261510/story.html (Accessed 11th December 2014).

Yore, L.D., Anderson, J.O. and Chiu, M. (2010) Moving PISA results into the policy arena: Perspectives on knowledge transfer for future considerations and preparations. *International Journal of Science and Mathematics Education.* 8, pp. 593-609.

Comparing U.S. States' Mathematics Results in PISA and Other International and National Student Assessments

Maria Stephens and Anindita Sen

For the first time in 2012, three U.S. states – Connecticut, Florida, and Massachusetts – participated in the OECD's Program for International Student Assessment (PISA) as individual entities in order to obtain an international benchmark of student performance. PISA measures students' reading, mathematics, and science literacy at 15 years of age, which is near the end of compulsory school in most of the participating countries. However, while this participation marked states' debut into PISA, which has been on going on a three-yearly basis since 2000, it was not the states' first foray into international student assessment. These three states, as well as others, have shown the same increasing interest in measuring their students against other students around the world that numerous countries themselves have shown (see Exhibit 1). For example, one of the three PISA-participant states was involved as early as 1995 in the IEA's administration of the Trends in International Mathematics and Science Study (TIMSS). U.S. states are now involved in all three major international student assessments, including the Progress in International Reading Literacy (PIRLS) Study.

This subnational participation in international assessments provides value nationally by contributing to a better understanding of the variation in national statistics and, for states, by providing a sense of the global comparative health of their education systems. However, one of the challenges in using and interpreting the international data, among the numerous other sources of information to which states have access to, is in understanding differing results across programs. This article thus focuses on the question: *What specific factors might explain differences in the United States' results on PISA 2012 and their results on other recent international and national assessments?* It describes

the results of a comparative analysis of four possible factors: (1) differences in overall content distribution of the items, (2) differences in relative strengths and weaknesses on content and cognitive subscales, (3) differences in sampling, and (4) differences in participating countries. It does not examine epistemological, ontological, or other methodological differences among the assessments. Data examined include the mathematics results from PISA 2012, TIMSS 2011 at the eighth grade, and the National Assessment of Educational Progress (NAEP) 2011 at the eighth grade.[1]

The article describes: (1) the educational contexts of the three PISA-participant states (Connecticut, Florida, and Massachusetts) and the assessment programs whose results are examined (PISA 2012, TIMSS 2011, and NAEP 2011) and, (2) differences in the key mathematics results for the three PISA-participant states on the three assessments, and (3) the results of the comparative analysis of factors potentially contributing to those differences.

Table 1: Countries' and jurisdictions' participation in PISA and other international student assessments: 1995-2012

Program	Year	No. of countries	No. of benchmarking jurisdictions[1]	No. of U.S. states among the benchmarking jurisdictions
PISA[2]	2012	64	4	3
	2009	75	5	0
	2006	57	0	0
	2003	41	0	0
	2000	43	0	0
TIMSS[3]	2011	63	14	9
	2007	55	8	2
	2003	49	4	1
	1999	39	27	13
	1995	43	6	6
PIRLS[4]	2011	48	9	1
	2006	41	0	0
	2001	35	0	0

1 "Benchmarking jurisdictions" refers to subnational entities that participate independently in an assessment – i.e., either representing an incomplete subset of a nation's subnational jurisdictions or those that finance their own participation in addition to the nation's participation. The OECD does not separately identify "benchmarking jurisdictions" because until 2009, no subnational jurisdictions participated independently. (China's two autonomous states of Hong Kong and Macao have participated since 2000 and 2003, respectively and are instead included in the country

1 Mathematics is examined because it was the focus of the 2012 PISA cycle.

count. Additionally, a number of federal countries have voluntarily oversampled in various years to provide for disaggregation within the national data and these cases are not counted as benchmarking jurisdictions.) For the purposes of this table, we have included the five subnational jurisdictions that were represented in PISA 2009 (one each from China, the United Arab Emirates [UAE], and Venezuela and two from India) and the four from 2012 (Shanghai-China and the three U.S. states). The IEA has historically treated the subnational jurisdictions of the Flemish and French communities of Belgium and the nations of the United Kingdom as individual education systems, on par with other national systems and these are included in the country count for TIMSS and PIRLS. However, they separately identify other subnational jurisdictions such as the various Emirates of the UAE, U.S. states, or Canadian provinces. This column does not include district or district consortia participation.

2 Counts include countries, jurisdictions, and states that administered a given year's assessment in the primary year or a follow-up wave (e.g., 2000 PISA in 2001/2 or 2009 PISA in 2010).

3 The counts include participants in 4th and/or 8th grade TIMSS.

4 Only more recently (2011) has the Progress in Reading Literacy Study (PIRLS) been opened for subnational participation. Florida was the U.S. state that participated in PIRLS 2011.

Background on State Education Systems and Assessment

Education in the United States is decentralized, with each state having responsibility for governing its own education system. These responsibilities including distributing federal and state funding, establishing policies (such as the duration of compulsory education, requirements for graduation, and minimum teacher qualifications), providing guidance regarding curriculum, conducting student assessments, and ensuring equal access to education for all eligible students. Often, some of these responsibilities — particularly those related to instruction — are further delegated to localities, which manage the operation of schools in their districts. While some aspects of education are very similar across states (e.g., the organization of schools), other characteristics (e.g., policies for compulsory education, demographics, teacher salaries) vary (see Exhibit 2, which provides a brief overview of education in the PISA-participant states).

The three PISA-participant states, as well as the other U.S. states, have access to a number of different macro-measures of student performance, and for the purposes of this article, we focus on those that can currently be compared across states, including the National Assessment of Educational Progress (NAEP) and the international assessments, PISA and TIMSS (see textbox for information and context on other macro-measures of student performance).

NAEP is the longest-standing measure of student performance for most states. The NAEP 4th- and 8th-grade assessments in reading and

mathematics, which are given every two years, are effectively required, as participation is a condition of receiving Title I funding, which is a primary financial resource (over $14 billion in 2014) for school districts and schools with high percentages of disadvantaged students (Federal Education Budget Project, 2014). The other NAEP assessments, including those at 12th grade and in other subjects, are voluntary. NAEP is designed to measure the knowledge and skills students have acquired in school on content determined through the collaborative input of a wide range of experts and participants from government, education, business, and public sectors in the United States. As the "nation's report card," NAEP is supposed to reflect what U.S. students should know and be able to do. For states, the benefit is the long trend line and the applicability to the U.S. context.

Table 2: Overview of Selected Education System Characteristics in 3 U.S. States: 2011-12

	Connecticut	Florida	Massachusetts
Governance			
Appoints the State Superintendent	State Board	Governor	State Board
Appoints the State Board	Governor	Governor	Governor
Structure			
Typical organization	Elementary education (Kindergarten through grade 5) Middle school (grade 6-8) High school (grade 9-12)		
Entrance age	Must be 5 by January 1 (of school year)	Must be 5 by September 1	Localities determine
Compulsory education	5-18	6-16	6-16
Demographics			
No. of districts	203	76	401
No. of schools	1,151	4,212	1,835
No. of students	554,457	2,668,156	953,369
No. of teachers	43,835	175,016	69,342
Student-teacher ratio	12.7	15.2	13.7
Percent of students FRPL[1]	34.5	56.9	34.2
Finance			
Total expenditure on public elementary and secondary education[2]	$9,094,036,286	$23,870,910,268	$13,649,965,365
Average annual salary of public elementary and secondary teachers 2011-12	$72,821	$46,232	$72,000

1 Reference year is 2010-11. FRPL is free and reduced price lunch, indicating students with lower socioeconomic resources.

2 Reference year is 2010-11.

Sources: NCES, 2014; ECS, 2014a; ECS, 2014b.

The international student assessments, PISA and TIMSS, are not required, though the Common Core initiative—described in the text box—has underscored the value of states' engaging in assessment in an international context. The Common Core initiative, which began around 2008 to increase rigor across state education systems, is both a result of and a driver of states' participation in international assessments. Since the 1995 administration of TIMSS, a total of 18 states have participated in at least one cycle of TIMSS, with 9 participating in multiple cycles and 9 participating in the most recent 2011 cycle. Additionally, as an indirect measure, states have looked to the estimates produced by the NAEP-TIMSS Linking Studies, the most recent of which used improved methodology to estimate TIMSS scores for each of the 50 states based on their NAEP scores and the NAEP and TIMSS results from the 9 states participating in both assessments in 2012 and 2011, respectively (NCES, 2013). This has been an important – if less reliable – source of information and significantly less costly than actual participation in international assessments.

Table 3: Overview of Selected Characteristics of Assessment Programs

	PISA 2012	TIMSS 2011	NAEP 2011
Frequency	Every 3 years	Every 4 years	Every 2 years
Target population	15 years old [1]	Grades 4 and 8	Grades 4, 8, and 12
No. of schools sampled [2]	~3000	~4,000	7,610
No. of students sampled [2]	~11,000	~37,000	175,200

1 In the United States, PISA's age-based national sampled included students mostly in grade 10 (71 percent in 2012), though some were in grade 11 (17 percent), grade 9 (12 percent), or other grades (less than 1 percent).
2 For TIMSS and NAEP, the numbers are for grade 8 only. For all three assessments, the numbers include state participants.
Sources: Provasnik et al., 2013; Mullis et al., 2012; NCES, 2012.

Evolving State Assessments and the Context for International Participation

All U.S. states also have state assessments. Some states have had assessment systems for decades, others initiated them in the 1990s with the passage of state accountability laws, and the rest developed or expanded them under the requirements of the No Child Left Behind Act (NCLB) 2001 (Chingos, 2012). NCLB required that states test all students in grades 3-8 and in one grade in high school in mathematics and reading. Prior to NCLB, only 13 states had assessment systems this extensive (Dantiz, 2001, as cited in Chingos, 2012). State assessments, however, are in the midst of another major change, as most states – with a boost from incentives from the federal level – have adopted the Common Core State Standards, which is an initiative that developed common standards in core academic subjects, and most are collaborating on the development of assessments of those standards that will replace their existing

systems. This will mean that, for the first time, there will be comparability in learning standards across states and in performance measures among at least some states. The lack of comparability and variable quality across states has been an often-cited weakness of NCLB in the past (e.g., Linn, Baker, and Betebenner, 2002).

The main purpose of the Common Core is to increase the rigor of standards and align them with the expectations of education institutions and employers so that students meeting the standards will be ready for college or a career.[2] A major driver of the Common Core was the states themselves – the initiative is managed by the Council of Chief State School Officers and the National Governors Association – and their expressed need for improved benchmarking – namely, "comparing outcomes to identify top performers or fast improvers, learning how they achieve great results and applying those lessons to improve...performance" (NGA, 2008, p. 9), with an explicit acknowledgement that the standards and benchmarks should have an international component. Thus not only should the standards be rigorous enough to allow U.S. students to compete in the global economy, states should measure performance in an international context (with implicit favour being given to PISA as the assessment of choice [Schneider, 2009]).

Forty-five states and the District of Columbia (including the three PISA-participant states) have adopted the Common Core standards in both English language arts and mathematics and an additional state in mathematics only. Two consortia, the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment (SBAC) Consortium, are the primary groups working on the new assessments that will roll out in the 2014-15 school year. Connecticut is signed on with SBAC, Massachusetts with PARCC, and Florida with another private provider.[2] One analysis has suggested that with a quality implementation of the Common Core in mathematics and well-designed assessment tasks particularly at the secondary level, U.S. students would be learning the kind of mathematics that would make them potentially more competitive in PISA (OECD, 2013).[3]

States have participated in PISA because of its targeting of students nearing the end of compulsory school and its focus on students' ability to apply the knowledge and skills they have learned cumulatively during their schooling, as well as in other contexts, for solving problems in a real-world context. States have participated in TIMSS, on the other hand,

2 Florida has contracted with American Institutes for Research (AIR) to develop its state assessments. AIR is the home organization of the authors of this article; the authors are in a separate division and independent of that project.

3 The referenced analysis classified the PISA 2012 items against the Common Core progression according to where they sit in the progression of standards up to high school level, the degree to which they represent attributes of modeling, and their modeling level. The analysis found a degree of commonality between the PISA and Common Core constructs leading the authors to conclude that "the high school curriculum in the United States will attend to modeling to a greater degree than has happened in the past...[and if] more students work on more and better modeling tasks than [they do] today, then one could reasonably expect PISA performance to improve" (p. 90).

because its grade-based target populations are similar to NAEP (grades 4 and 8) and it also similarly focuses on school achievement. Both the national and two international assessments collect data on mathematics and science performance, though the sampling requirements and other features differ (see Exhibit 3).

U.S. States' Mathematics Results from PISA, TIMSS, and NAEP

While each of the sources of student performance data provides valuable input for U.S. states, interpretation of results across the multiple measures requires careful consideration. (Again, since the state assessments aligned with the Common Core are not fully in place yet, we focus here on the data available from PISA and how that aligns with data available from TIMSS and NAEP.)

On average, students in the United States performed below the OECD average in mathematics literacy, scoring 481 points compared to the OECD mean of 494 in 2012 (see Exhibit 4 and Kelly et al., 2013). This masks variation among the states, however. Students in Connecticut scored an average of 506 points, which was above the U.S. mean though statistically comparable to the OECD mean. Just 12 of the 68 total participating education systems scored higher than Connecticut and its scores were comparable to those of students in 15 other systems, including the United States' partners in the G-8 Canada, France, Germany, and the United Kingdom. Students in Massachusetts scored an average of 514 points, which was statistically comparable to Connecticut's mean but above both the U.S. and OECD means. Nine education systems outperformed Massachusetts and its scores bested an additional six education systems than did Connecticut's. In contrast, students in Florida scored 467 points on average, which was lower than both the U.S. and OECD means. Florida's mean score was below that of 38 education systems and statistically comparable to a set of five education systems—Lithuania, Sweden, Hungary, Croatia, and Israel—that were outperformed by the other two PISA-participant states. These findings were not necessarily surprising as the two Northeastern states are typically above average in NAEP and Florida is typically below-average, as they were in 2011.

Looking across the assessments highlights some differences and generates interesting questions. Connecticut performed above the U.S. mean in mathematics in PISA 2012 and eighth-grade NAEP 2011, and above the international mean in TIMSS 2011 though similar to the OECD mean in PISA 2012. (It should be noted that the OECD mean in PISA is based only on the scores of the participating OECD education systems whereas

the international mean in TIMSS is based on all the participating TIMSS education systems, which is a much more diverse group in terms of student outcomes.) However, despite the differences in performance relative to the international means, Connecticut appears to have a greater advantage in PISA than in TIMSS (and NAEP), based just on distance from the U.S. mean. What might account for this advantage?

Table 4: Mathematics performance of U.S. 15-year-olds in PISA and eighth-grade students in TIMSS and NAEP: 2011 and 2012

	PISA 2012 (15-year-olds)			TIMSS 2011 (Grade 8)			NAEP 2011 (Grade 8)	
	Mean score	Relative to U.S.	Relative to OECD	Mean score	Relative to U.S.	Relative to Int'l	Mean score	Relative to U.S.
Connecticut	506	.	.	518	.	.	287	+
Florida	467	.	.	513	.	.	278	.
Massachusetts	514	.	.	561	.	.	299	+
United States	481	–	.	509	–	.	285	–
Across countries	494	.	.	500	.	.	.	.
Range across 50 states	.	.	.	466 (Ala.) – 561 (Mass.)[1]			267 (DC) -299 (Mass.)	
Range across countries	368 (Peru) – 613 (Shanghai-China)			331 (Ghana) – 613 (Korea, Rep. of)			[1]	

Note: PISA measures mathematics literacy, or the application of mathematics for solving real-world problems. TIMSS and NAEP focus more exclusively on school-based mathematics.

1 The range is based on scores estimated in the NAEP-TIMSS Linking Study; results for the three PISA states, however, are actual TIMSS results as they also participated in TIMSS 2011.

. Not applicable

+ Significantly higher than reference at the .05 level.

– Significantly lower than reference at the .05 level.

Not significantly different than reference at the .05 level.

Sources: Kelly et al., 2013; Mullis et al., 2012; NCES, 2012; and NCES, 2013.

Massachusetts also performed above the U.S. mean in mathematics on all three assessments, as well as above the respective international means for PISA 2012 and TIMSS 2011. Based on distance from the U.S. mean, however, Massachusetts appears to have a greater advantage in TIMSS (and NAEP) than PISA. Again, what might account for this particular advantage?

Finally, Florida performed lower than the U.S. mean in mathematics in PISA 2012 and eighth-grade NAEP 2011, but similar to the U.S. mean in TIMSS 2011. On the international assessments, despite being lower than the OECD mean in PISA 2012, Florida is above the TIMSS 2011 international mean. How would Florida's relative standing change if the groups of education systems participating in PISA and TIMSS were comparable? What are some possible explanations for Florida's weaker-than-average performance?

Analysis of Differences in Results

A first analysis to explore these questions is to examine the similarities and differences in terms of item content, which has been collected through studies comparing the various international assessments with each other and with NAEP.[4] Generally speaking, these studies have shown that, overall, there are more similarities between NAEP and TIMSS than between NAEP and PISA, as might be expected given the former two programs' focus on curriculum-based achievement and the latter's on literacy (Provasnik et al., 2013; AIR, 2013). For example, PISA differs from TIMSS and NAEP in terms of the distribution of test items across content areas: PISA 2012 had a larger percentage of items that would be considered data analysis, probability and statistics items on the NAEP framework than did NAEP 2011/2013 or TIMSS 2011, whereas it had a smaller percentage of items classified as algebra (see Exhibit 5).[5] Additionally, the most recent comparison study identified several topics covered by the NAEP 2013 item pool that were not covered by the PISA 2012 item pool – i.e., that were unique to NAEP – including: estimation; mathematical reasoning using numbers; position, direction, and coordinate geometry; mathematical reasoning in geometry; measurement in triangles; experiments and samples; mathematical reasoning with data; and mathematical reasoning in algebra (AIR, 2013). In terms of item complexity, PISA 2012 had a greater percentage of items classified as "moderate" on the NAEP framework than did NAEP 2013, and a smaller percentage classified as "low" (data not shown, AIR, 2013).

4 See http://nces.ed.gov/surveys/international/cross-study-comparisons.asp for a listing of these studies through 2013.

5 This is based on results from two studies: one (Lin, Darling, and Dodson, 2013) that compared the NAEP 2011 and TIMSS 2011 grade 8 mathematics items (among other elements) and another (AIR, 2013) that compared the NAEP 2013 grade 8 and PISA 2012 items (among other elements). Though different expert panels undertook the studies, the distribution of NAEP grade 8 mathematics items across content areas was assessed similarly by the two groups.

Table 5: Distribution of items across NAEP mathematics content areas

Content Areas in the NAEP Framework	PISA 2012[1]	NAEP 2013 Grade 8	NAEP 2011 Grade 8	TIMSS 2011
Number properties and operations	33%	19%	17%	28%
Geometry	14%	17%	17%	9%
Measurement	16%	19%	19%	12%
Data analysis, probability and statistics	27%	15%	14%	18%
Algebra	11%	30%	32%	34%

1 This is based on the 64 (of 85) PISA items that were classified to the NAEP grade 8 framework.
Sources: Provasnik et al., 2013; AIR, 2013.

So, theoretically if students in Connecticut – where there appears to be a relative advantage in PISA – have had greater exposure to data analysis, probability, and statistics items or items of similar nature or complexity to PISA items, this might contribute to their relatively strong performance in PISA. On the other hand, if students in Massachusetts – where there is a relative advantage in TIMSS – have had a strong focus on algebra this could partly explain the excellence in TIMSS and NAEP. This could be explored by examining the state standards and assessments in place.

A second analysis examines states' scores on the mathematics subscales, which in PISA 2012 included three processes (employ, formulate, and interpret) and four content categories (space and shape, change and relationships, quantity, and uncertainty) to determine if states' relative strengths and weaknesses align with relative areas of emphasis or de-emphasis in the various assessments. For example, Connecticut was comparatively strong in items requiring interpretation, of which there were a larger percentage in PISA 2012 than in NAEP 2013 (see Exhibit 6). Items in the interpretation category were a relative strength for all states, however. In terms of the content subscales, there were again similar patterns among the PISA-participant states, with change and relationships (i.e., algebra) and uncertainty (i.e., probability and statistics) as relative strengths and quantity and space and shape as relative weaknesses. It is difficult to relate these results to item distributions in NAEP, however, because in the comparison study on which the data are based, a high percentage of NAEP 2013 items were found not to fit the PISA framework.

A third analysis relates to sampling. As PISA uses an age-based sample, sampled students may come from various grades, which is a distinction from TIMSS and NAEP. This feature of PISA is in keeping with

Table 6: Mathematics performance and percentage distribution of items by PISA process and content subscales

	Employ	Process subscales			Content subscales			
		Formulate	Interpret	Space and shape	Change and relationships	Quantity	Uncertainty	
Mean score	Connecticut	502	504	515	487	515	502	512
	Florida	466	458	475	446	476	458	475
	Massachusetts	509	512	524	498	518	506	523
	United States	480	476	490	463	488	478	488
	OECD	493	492	497	490	493	495	495
Percentage	PISA 2012	44%	32%	25%	25%	25%	26%	25%
	NAEP 2015[1]	66%	25%	9%	7%	14%	6%	7%

[1] The percentages for NAEP items in the content categories will not sum to 100 because 66 percent of the NAEP eighth-grade items were found not to fit the PISA framework.

Source: AIR, 2015 and PISA International Data Explorer (http://nces.ed.gov/surveys/international/ide/).

its goal to measure the outcomes of learning, rather than schooling *per se* and provides a neutral comparison point internationally. Intra-nationally, in the case of federal systems with variation in local education policy, this can be a source of some differences. For example, analysing the grade distribution of the students who took PISA in 2012 shows that Connecticut had a larger percentage of students in the 11th grade and smaller percentages in the 9th and 10th grades than Florida, Massachusetts, or the United States overall (see Exhibit 7). Conversely, Florida had a larger percentage of students in the 9th grade and smaller percentages in the upper grades than the other systems. In other words, a larger percentage of Connecticut's students were exposed to an additional year of schooling than were U.S. students on average or in Massachusetts or Florida. And a larger percentage of Florida's students had not yet been exposed to 10th- or 11th-grade mathematics than had students in the other systems. This is due to differences in policies on school entry and in grade retention practices. For example, Connecticut has one of the youngest kindergarten entry ages in the United States, allowing students to enrol at 4 years old as long as they will be 5 years old by mid-school year (e.g., January 1) and requiring enrolment at 5 (ECS, 2014; see also Exhibit 2). Other states more typically have cut-offs early in the school year, requiring that students be 5 years old, e.g., by September 1 and not requiring enrolment until 6, as in Florida and Massachusetts. What may then account for Florida's higher

rate of 9th-grade PISA participants are generally higher early grade retention rates than in the other two states (Warren and Sariba, 2012). Of the analyses described, the sampling explanations appear to have the strongest explanatory potential.

Table 7: Distribution of PISA participants by grade: 2012

	7th grade	8th grade	9th grade	10th grade	11th grade	12th grade
Connecticut	*	*	7	59	34	*
Florida	*	*	21	67	12	*
Massachusetts	*	*	1	82	17	*
United States	*	*	12	71	17	*

* Reporting standards not met.
Note: Results for Connecticut, Florida, and Massachusetts are for public school students only. Detail may not sum to totals because of rounding. Some apparent differences between estimates may not be statistically significant.
Source: PISA International Data Explorer (http://nces.ed.gov/surveys/international/ide/).

A final analysis questions the differing country populations in PISA and TIMSS: how would states' standings relative to the OECD/international means change if the assessments included the same group of countries? Restricting the countries in the analyses to only those that participated in both PISA 2012 and TIMSS 2011 at eighth grade,[6] both the OECD and international averages drop. So, while this brings Florida's mean score closer to the PISA OECD mean score (though still statistically significantly below it), it further distances the state's mean score in a positive direction from the TIMSS international mean – essentially, leaving the relative standings unchanged.

Conclusion

U.S. states' participation in international assessments shows one source of variation in national statistics and also allows states to benchmark themselves to international standards, as has been shown to be an increasing interest over at least the last decade. However, given that states also have access to national assessment data, as well as their own state data and, in some cases, two sources of international data, making sense of the results can be challenging. Analyses described in this paper suggest that opportunity to learn may be an important factor in differing results among assessments – with the amount of schooling related to states' PISA per-

6 This represents 28 countries, with the only difference being the participation of all nations of the United Kingdom in PISA versus only England in TIMSS. The referenced analysis is based on data (not shown) obtained from the PISA International Data Explorer.

formance. A next frontier for state participants in international student assessments will be in how they, and their localities, may try to extend the use of data beyond the core benchmarking function to absorb lessons from international partners and inform education policy.

References

American Institutes for Research (2013) *A comparison study of the Program for International Student Assessment (PISA) 2012 and the National Assessment of Educational Progress (NAEP) 2013 mathematics assessments*. Washington, D.C.: Author.

Chingos, M.M. (2012) *Strength in numbers: State spending on K-12 assessment systems*. Washington, D.C.: Brown Center on Education Policy, The Brookings Institution. Available from: http://www.brookings.edu/~/media/research/files/reports/2012/11/29%20cost%20of%20assessment%20chingos/11_assessment_chingos_final.pdf (Accessed 7th July 2014).

Danitz, T. (2001, 27th February) Special report: States pay $400 Million for tests in 2001. *Stateline*. Pew Center on the States.

Education Commission on the States (2014) Early learning: Kindergarten online database [data set]. Available from: http://www.ecs.org/html/educationissues/kindergarten/kdb_intro_sf.asp (Accessed 7th July 2014).

Education Commission on the States (2014) What Governors Need to Know: Highlights of State Education Systems. Available from: http://www.ecs.org/clearinghouse/85/69/8569.pdf (Accessed 7th July 2014).

Federal Education Budget Project (2014) Available from: http://febp.newamerica.net/background-analysis/no-child-left-behind-funding (Accessed 7th July 2014).

Kelly, D., Xie, H., Nord, C.W., Jenkins, F., Chan, J.Y., and Katsberg, D. (2013) *Performance of U.S. 15-year-old students in mathematics, science, and reading literacy in an international context: First look at PISA 2012* (NCES 2014-024). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Available from: http://nces.ed.gov/pubs2014/2014024rev.pdf (Accessed 7th July 2014).

Linn, R.L., Baker, E.L., and Betebenner, D.W. (2002) Accountability systems: Implications of the requirements of the No Child Left Behind Act of 2001. *Education Researcher. 31 (6)*, pp. 3-16.

Provasnik, S., Lin, C., Darling, D., and Dodson, J. (2013) *A comparison of the 2011 Trends in International Mathematics and Science Study*

(TIMSS) assessment items and the 2011 National Assessment of Educational Progress (NAEP) frameworks. Washington, D.C.: AVAR. Available from: http://nces.ed.gov/nationsreportcard/subject/about/pdf/naep_timss_comparison_items.pdf (Accessed 7th July 2014).

Mullis, I.V.S, Martin, M.O., Foy, P., and Arora, A. (2012) *TIMSS 2011 international results in mathematics.* Boston, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

National Centre for Education Statistics (2012) *Mathematics 2011: National Assessment of Educational Progress at grades 4 and 8, NCES 2012-458.* Washington, D.C.: NCES, Institute for Education Sciences, U.S. Department of Education.

National Centre for Education Statistics (2014) Tables and Figures. Available from: http://nces.ed.gov/quicktables/index.asp. (Accessed 29th September 2014).

National Centre for Education Statistics (NCES) (2013). *U.S. states in a global context: Results from the 2011 NAEP-TIMSS linking study, NCES 2013-460.* Washington, D.C.: NCES, Institute for Education Sciences, U.S. Department of Education. Available from: http://nces.ed.gov/nationsreportcard/subject/publications/studies/pdf/2013460.pdf.

National Governors Association (2008) *Benchmarking for success: Ensuring U.S. students receive a world-class education.* Washington, D.C.: NGA. Available from: http://www.corestandards.org/assets/0812BENCHMARKING.pdf (Accessed 7th July 2014).

OECD (2013) *Strong performers and successful reformers in education – Lessons from PISA 2012 for the United States.* Paris: OECD. Available from: http://www.oecd.org/pisa/keyfindings/PISA2012-US-CHAP4.pdf (Accessed 7th July 2014).

Schneider, M. (2009) The international PISA test. *Education Next, 9, 4.* Available from: http://educationnext.org/the-international-pisa-test/ (Accessed 19th June 2014).

Warren, J.R., and Saliba, J. (2012, May). *Public school grade retention rates in the United States: Estimates by state, grade, year, and race/ethnicity.* Paper presented at the Population Association of America Annual Meeting, San Francisco, CA.

The Predictive Power of Attribution Styles for PISA 2012 Achievement: International and National Perspective

Ana Kozina and Ana Mlekuž

Introduction

In the paper, Weiner's attribution theory is used as a framework in explaining the differences between high and low achieving students in PISA 2012 study for international and national analyses.

The reasons people give for why they succeeded or failed a task are called attributions. (Heider, 1958, in Nokelainen, Tirri and Merenti-Valimaki, 2007). Furthermore, attribution theory has been widely recognized as a significant contributor in achievement explaining models (Stroud and Reynolds, 2009). According to Dembo and Eaton (1996, in Stroud and Reynolds, 2009), motivation is constructed from three internal factors, one of them being the students' attributions for success and failure (the other two are: the importance placed on the task and the emotional process associated with the learning process). Weiner (1985; 2010) defined attributions more precisely. He distinguished attributions on three dimensions: locus (whether the cause is internal or external), controllability (whether the cause can be subjected to volitional influence) and stability (whether the cause is stable or varies over time). He also identified four common attributions that differ on these dimensions: effort (internal, controllable and unstable), ability (internal, uncontrollable and stable), task difficulty (external, uncontrollable, stable) and luck (external, uncontrollable, and unstable). Additionally, attribution constructs can be classified into three groups: attribution appraisals (explanations assessed following actual or manipulated success or failure in performing a specific task), attribution beliefs (domain specific or domain general beliefs about the causes of success or failure), attribution styles (generalized, stereo-

typical patterns of attributions and dispositional beliefs) (Dai, Moon and Feldhusen, 1998, in Nokelainen et al., 2007).

The specific attributions that students make affect their expectancy for future performance, persistence in similar tasks, emotional responses, which tasks they choose, and self-efficacy, which is an important characteristic for educational setting (Demo and Eaton, 1996, in Stroud and Reynolds, 2009). Students with an internal locus of control believe that events in life are controlled by their own actions, whereas those with an external locus of control attribute the outcomes of events to outside factors such as luck. In general, people with an external locus of control appear to be prone to a variety of symptoms of stress including emotional distress, job dissatisfaction, burn-out and low self-esteem (Matthews, Deary and Whiteman, 2009). On one hand, students with attributions showing the internal locus of control (e.g. effort) will work harder to improve themselves in school. In addition to this, those students who attribute their success or failure to external factors (e.g. parents, friends, teachers...) tend not to invest more time in learning.

The motivational path of causal attribution begins with the interpretation of the event (in our case the mathematics achievement) as success or failure. Following the initial reaction of happiness or sadness, individuals search the reason why this specific outcome has occurred. In the achievement domain, successes and failures are often attributed to an ability factor, an effort factor, the difficulty of the task, luck, mood and help or hindrance from others. When explaining achievement results, individuals attach the most importance to their perceived competences and how hard they tried. The attribution theory proposes that people spontaneously engage in such causal thinking in their everyday lives (Graham and Williams, 2009).

Studies broadly investigated the relationship between attribution styles and academic achievement (Gibb et al., 2002) stating a significant relationship and significant predictive value of the locus of control for academic achievement (Gibb et al., 2002; Philips and Gully, 1997), study time and effort (Shell and Husman, 2008). For instance McClure, Meyer, Garisch, Fischer, Weir and Walkey (2011) examined the relationship between attributions for success and failure and academic achievement among students aged 14 and 15 years (as in PISA study). They also measured motivation orientations and cultural differences; therefore European, Asian, Maori and Pacific participants were included in the research. The measure assessed attributions (causes for their best and worst performance only), motivation orientation (doing my best and doing just enough scales), demographic data and achievement data. The results firstly confirmed the self-serving bias, which was already proven in many pre-

vious studies (e.g. Bong, 2004; Vispoel and Austin, 1995). Students show a self-serving pattern of attributing their highest marks to effort and ability more than their lowest marks, which are mostly attributed to task difficulty. Students who attributed their best marks to internal factors of ability and effort attained higher achievement. On the other hand, students who attributed their best marks to luck, family and friends gained lower achievement scores. Moreover, attributions for their worst marks were also important. Students who attributed their worst marks to ability, effort, high task difficulty and the influence of teachers gained higher achievement scores, whereas students who attributed their worst marks to family and friends gained lower achievement scores. In addition, the regression analyses showed that the students' motivation orientation and attributions is a significant predictor of achievement, accounting for 38 % of the students' achievement scores. Among attributions the strongest positive predictor was attributing the best marks to effort and the worst marks to lack of effort and to the influence or characteristics of the teacher, while the main negative predictors were attributing the best or worst marks to family and friends and attributing the best marks to luck.

Similar patterns were established in primary school students. Khodayarifard, Brinthaupt and Anshel (2010) examined the relationships between academic achievement and the child's and the parent's attribution styles in primary school students and their parents. Regarding the connection between attributions and academic achievement, the results were consistent with previous research (Carr et.al. 1991; Stipek and Hoffman, 1988). Students who did not perform well academically tended to show a more negative attributional style (attributing negative events to more stable and uncontrollable causes).

Longitudinal effects were tested in Liu, Cheng, Chen and Wu (2009) study. They examined the longitudinal effect of educational expectations and achievement attributions on adolescents' academic achievement (secondary school students). The results show that high educational expectations and attribution to effort (controllable, unstable attribution) have a positive effect on learning growth rate, while attributions to others have a negative effect on the learning growth rate. Furthermore, as already proven in previous research (e.g. Georgiou, 1999), attributions of achievements to effort are positively related to actual achievements, whereas attributions to others are negatively related to achievement. The pattern of perceived control is associated with better self-regulation, knowledge building, question asking, study use and effort (Shell and Husman, 2008). The study showed that such attributional patterns influence the long-term academic development of adolescents (Schunk, 1992).

The relationship between attribution styles and academic achievement can be explained using the concept of self-regulation. According to social-cognitive theory, self-regulation is dependent on the situation and it is not stable. Based on this assumption, Zimmerman (2000) describes self-regulation as cyclical with three phases containing sub processes: forethought (task analyses and self-motivation beliefs), performance (self-control and self-observation), and self-reflection (self-judgement (e.g. self-evaluation and causal attribution) and self-reaction (e.g. self-satisfaction)). According to their performance in each of these domains, learners have been described as skilled or unskilled learners (Stroud & Reynolds, 2009). Attributions are a part of the final stage. Self-reflection begins with self-judgement (individual comparisons of information gained through self-monitoring to extrinsic standards or goals). An individual is motivated to have fast and accurate feedback on his or hers performance as compared to others. Self-judgement leads to attribution interpretations where the learner interprets the reasons for success and failure. Attribution interpretations can lead to positive self-reactions. The individual might interpret their failure as the result of too little effort and then increase his or hers efforts. On the other hand, if they interpret their failure as a lack of ability the reaction is likely to be decreased in learning behaviour. Attribution interpretations reveal the possible reasons for learning mistakes and help the learner to find the most appropriate learning strategies. Additionally, they also promote adaptation and self-regulation, which eventually leads to a more positive self-image and enhance intrinsic interest in the task (Nokelainen et al., 2007). Ellström (2001, in Nokelainen et al., 2007) goes even beyond that stating that attributions for success and failure affect potential competence.

Attribution style has been shown in some studies to alter according to the context (Sarafino, 2006, in Graham and Williams, 2009). Therefore the focus of this paper is mainly on the educational setting and on mathematical achievement. The paper concentrates specifically on PISA 2012 results and the predictive value of attribution styles on PISA 2012 mathematics achievement. PISA measures attribution styles in the context of the students' drive and motivation in the form of separate questions in the students' background questionnaire. PISA measures drive and motivation using four concepts: perseverance (constructed index based on the students' responses about their willingness to work on problems that are difficult, even when they encounter problems), openness to problem solving (constructed index based on the students' responses about their willingness to engage with problems), locus of control/attribution style (constructed index based on the students' responses about whether

they attribute failure in mathematics test to themselves or to others; and the students responses about whether they strongly agree that success in mathematics and school depends on whether they put in enough effort) and motivation to learn mathematics – intrinsic and instrumental (constructed indices based on the students' responses about whether they enjoy mathematics and work hard in mathematics because they enjoy the subject, and whether they believe mathematics is important for their future studies and careers) (OECD, 2013b). In line with the attribution theory, PISA measures attributions on all three dimensions (locus, control, stability). Exposing individuals to academic success or failure and then asking them to report about their feelings and thoughts can measure attribution styles. The other possibility is to design a set of items where individuals imagine success or failure and then self-report what their most likely thoughts would be as is the case in the PISA study.

The present study aims to:

(1) Identify the attribution for success question set structure on an international level: All constructs that measure drive and motivation in PISA are developed in a form of indices on an international level except the question set measuring attribution for success (the students' responses about whether they strongly agree that success in mathematics and school depends on whether they put in enough effort) therefore the first aim of this study is to analyse the structure of this question set at the international level in order to construct an index that could be used as predictors in second aim of the study.

(2) Analyse predictive power of the attribution for success in mathematics for mathematics achievement on an international and national (Slovenia) level. The second aim of the study therefore is to use the newly developed index (indices) as a predicting variable in a regression model for mathematics achievement on an international level. Our basic assumption in line with the theoretical framework is that an internal locus of control predicts higher achievement on an international and national level. To test the generalizability of our findings we will use the same regression model on an international level (PISA 2012 international data base) on national level (Slovene PISA 2012 data base) and additionally in selected EU member states with different average mathematics achievement score. The choice was made based on average students' mathematics achievement score (as presented in international reports), where Netherlands and Estonia are the EU member states with the highest achievement score and Bulgaria and Romania are the EU member states with the lowest

achievement score[1]. In addition to the international data results and the results for Slovenia and four other countries' results will be analysed in detail. The goal is to test whether the same predictions can be made in high and in low achieving countries. Since attribution styles are under the strong influence of culture (e.g. western cultures valuing ability more and eastern countries valuing effort more) (Nokelainen et al., 2007) we have chosen EU member states for the comparisons.

Method

Participants

In the analyses, a PISA international sample is used. PISA samples students aged between 15 and 16 years, disregarding the grade levels or type of institution in which they are enrolled and regardless of whether they are in full-time or part-time education. Therefore, the average age of students included in the survey is 15 years and 9 months (OECD, 2014). Most countries included in PISA used a two-stage stratified sampling design, which means that the sampling was conducted in two stages. The first stage consisted of sampling individual schools, where 15-year-old students might be enrolled. A minimum of 150 schools per country were sampled. The second stage of the sampling process consisted of sampling 15 year-old students at the selected schools. Approximately 35 15-year-old students were sampled per school with equal probability, however each country then chooses its own modified sampling design (OECD, 2014). With these sampling procedures the representativeness of the selected test population for each educational system was ensured.

PISA 2012 focused on mathematical literacy. There were approximately 510 000 students from 65 countries included in the survey. For the purposes of this article data from the Form B Questionnaire and Slovene, Bulgarian, Romanian, Estonian and Dutch data sets are used (N=309 104). Each student answered a cognitive test and a background questionnaire. PISA 2012 introduced a new rotation design for the student questionnaire, which is similar to the cognitive items design. Items are combined in packages, which are distributed over a number of different booklets. Each student is assigned one of these booklets and therefore receives a limited number of items, whereas all booklets together cover a larger pool of items from different scopes (OECD, 2013c).

1 Even though Cyprus was the EU member state with the second lowest mathematics achievement score, it was not included in the analysis since there were no available data for this country in the international database.

Table 1: Samples characteristics

	N	Gender (%)		Average achievement score of all students included in PISA 2012 (s. e.)
		Female (s. e.)	Male (s. e.)	
Slovenia	3 706	49 (0.9)	51 (0.9)	501 (1.2)
Netherlands	2 757	49 (0.8)	51 (0.8)	523 (3.5)
Estonia	3 127	51 (0.8)	49 (0.8)	521 (2.1)
Romania	5 314	51 (1.5)	49 (1.5)	445 (3.8)
Bulgaria	3 299	48 (1.9)	52 (1.9)	439 (4.1)

Note: All the data presented in this table are calculated using only the data for students who answered question ST43 (attribution for success) in Student Questionnaire.

In Slovenia 3 706 students were included (49% female and 51% male). The average mathematics achievement score for Slovenia is 501, whereas for Netherlands, which is the EU member state with the highest score, the average students' achievement score is 523 on the other hand for Bulgaria, the EU member state with the lowest score, the average students' achievement score is 439.

For the data analysis, two programmes were used as follows: SPSS for structures analysis and IDB Analyser for regression analysis.

Instruments

Background Questionnaires

In PISA 2012, students completed a 30-minute student questionnaire, which included questions on their background, attitudes toward mathematics and on their learning strategies (OECD, 2013c). These questions are of vital importance for the analyses of the results. In detail, the questionnaire includes:

- student and their family background (including their economic, social and cultural capital),
- aspects of the students' lives (their attitudes towards learning, their habits and life inside school, their family environment),
- aspects of learning and instruction in mathematics, including the students' interest, motivation and engagement (OECD, 2013c).

Cognitive Tests

PISA 2012 was composed of a paper-based assessment of the students' mathematics, science and reading literacy and a computer-based assessment of problem solving (NCES, 2014a). All PISA 2012 cognitive items were organized in clusters. The main competency tested in PISA 2012 was

mathematical literacy. There were two possibilities to assess the mathematical literacy for countries. The first possibility was a set of 13 booklets, which included items distributed across a range of difficulty. Out of 7 mathematical clusters, 4 were included in these booklets according to a rotated test design. The booklets also included 3 reading clusters and 3 science clusters. Moreover, in each booklet there was at least one mathematical cluster. Regardless of a specific countries' choice, the performance of students in all participating countries is represented on a common mathematical literacy scale (OECD, 2013a).

Included Variables

Achievement scores for mathematics (Plausible values)

Each student had a different subset of items in their booklet; therefore scaling techniques were used to establish a common scale for all students. In PISA 2012, item response theory (IRT) was used to estimate average scores for measured competencies (mathematics included). This theory identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of the students' proficiency in answering other questions (NCES, 2014b).

Since each student completed only a subset of items, the students' scores are estimated as plausible values.[2] For each student five plausible values are estimated. These values represent the distribution of potential scores for all students in the population with similar characteristics and identical patterns of item response (NCES, 2014b).

Attribution for success in mathematics

The attribution for success in mathematics (internal and external) is measured with a set of questions (Question ST43). The question set measures the students' perceived control over their success in mathematics. This question examined the students' agreement with six statements about their mathematics lessons. Students had to evaluate whether they strongly agree (1)[3], agree (2), disagree (3) or strongly disagree (4) with the following statements: *If I put in enough effort I can succeed in mathematics; Whether or not I do well in mathematics is completely up to me; Family demands or other problems prevent me from putting a lot of time into my mathematics work; If I had different teachers, I would try harder in mathematics; If I wanted to, I could do well in mathematics; I do badly in mathematics whether or not I study for my exams* (OECD, 2013a).

2 More information on plausible values can be found in PISA Analysis Manual: http://browse.oecdbookshop.org/oecd/pdfs/free/9809031e.pdf.

3 Values in brackets are values of the variable entered in the database.

Attribution for failure in mathematics (perceived self-responsibility for failing mathematics)

Perceived self-responsibility for failing mathematics (FAILMAT) is an index constructed from students' responses to a set of questions from the background questionnaire. The questions examined the following situation: "*Suppose you are a student in the following situation: each week, your mathematics teacher gives you a short quiz. Recently you have done badly in these quizzes. Today you are trying to figure out why.*" Then followed six sentences that students had to evaluate whether they are very likely (1)[4], likely (2), slightly likely (3) or not at all likely (4) to have the following thoughts or feelings about this situation. The sentences describing the thought or feelings were as follows: *I'm not very good at solving mathematics problems; My teacher did not explain the concepts well this week; This week I made bad guesses on the quiz; Sometimes the course material is too hard; The teacher did not get students interested in the material; Sometimes I'm just unlucky* (OECD, 2013a).

Results

Structures Analyses

In order to define the underlying structure of the question set measuring the attribution for success in mathematics, the correlation matrix of the question set was subjected to factor analyses (method: principal axis factoring) on the international database. The preliminary test showed the data was suitable for this kind of analyses (KMO=0.670; Bartlett's Test of Sphericity (15) = 281867.271; $p<.001$).

Table 2: Total variance explained

Factor	Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %
Internal locus of control	2.065	34.411	34.411
External locus of control	1.490	24.841	59.252

The Kaiser-Guttmann criteria (Eigenvalues over 1) revealed two factors explaining 59.25 % of the question set total variance (Table 2). Based on the factor loading (Table 3) we named the two factors: *internal locus of control* and *external locus of control*. The first factor explains 34.411 % of variance and includes internal attributions for success in form of effort (internal, unstable, controllable) and own responsibility for success. The second factor explains 24.841 % of variance and includes external attributions for success in forms of external causes (external (e.g. teachers, family), stable, uncontrollable).

4 Values in brackets are values of the variable entered in the database.

Table 3: Factor loadings of attribution for success in mathematics question set

	Factor	
	Internal locus of control	External locus of control
If I put in enough effort, I can succeed in mathematics	.757	-.075
Whether or not I do well in mathematics is completely up to me	.634	-.038
Family demands or other problems prevent me from putting a lot of time into my mathematics work	-.069	.546
If I had different teachers, I would try harder in mathematics	.006	.514
If I wanted to, I could do well in mathematics	.671	.072
I do badly in mathematics whether or not I study for my exams	-.253	.529

Regression Analyses

For the analysis of the relationship between attribution for success in mathematics and the students' mathematics achievement, regression analysis was used. The regression analyses are at the first stage of the analyses conducted on an international level, and further on also on a national (Slovene) level followed by international comparisons. We used two stages of multiple regression analyses. In the first stage, only attributions for success in mathematics indices were entered in the model. Furthermore, in the second stage, attributions for failure in mathematics index (perceived self-responsibility for failing mathematics) were added to the model on national and international level.

A multicollinearity assumption of predictors in the model was tested with correlation analyses. All indices (*internal locus of control, external locus of control and perceived self-responsibility for failing mathematics*) statistically significantly correlate with each other, either weakly or moderately ($0.02< r <0.34$). Additionally VIF[5] were significantly below 10 ($1.08<VIF<1.15$).

As it can be seen from the Table 4, international data shows attribution for success in mathematics (internal and external locus oficontrol) as significant predictors of mathematics achievement in PISA 2012. The in-

5 In multiple regression, the variance inflation factor (VIF) is used as an indicator of multicollinearity. Computationally, it is defined as the reciprocal of tolerance: $1/(1-R^2)$. Various recommendations for acceptable levels of VIF have been published in the literature. Perhaps most commonly, a value of 10 has been recommended as the maximum level of VIF (Field, 2000).

Table 4: Predictive power of attribution for success in mathematics (internal and external locus of control) for mathematics achievement in PISA 2012 – international and national level

International results	b (s.e.)	β (s.e.)	R^{2*} (s.e.)
constant	478.18* (2.59)		
Internal locus of control	-16.56* (2.29)	-0.16* (0.02)	
External locus of control	32.84* (2.58)	0.26* (0.02)	0.11 (0.03)
Slovenia			
constant	503.79* (1.6)		
Internal locus of control	-15.27* (2.11)	-0.14* (0.02)	
External locus of control	28.29* (2.73)	0.23* (0.02)	0.08 (0.01)
Netherlands			
Constant	529.78* (3.77)		
Internal locus of control	-14.99* (2.22)	-0.16* (0.02)	
External locus of control	36.26* (4.43)	0.26* (0.03)	0.11 (0.02)
Estonia			
constant	516.20* (2.06)		
Internal locus of control	-18.13* (2.18)	-0.18* (0.02)	
External locus of control	37.82* (2.37)	0.31* (0.02)	0.15 (0.01)
Romania			
constant	454.38* (3.86)		
Internal locus of control	-10.46* (2.59)	-0.11* (0.03)	
External locus of control	29.00* (3.13)	0.28* (0.03)	0.09 (0.02)
Bulgaria			
constant	444.41* (3.46)		
Internal locus of control	-12.71* (2.16)	-0.12* (0.02)	
External locus of control	42.96* (2.59)	0.36* (0.02)	0.14 (0.02)

Notes: The data are weighted with Final Student Weight. R^{2*} is adjusted R^2. All the data presented in this table are calculated using only the data for students who answered question ST43 (attribution for success) in Student Questionnaire. Statistically significant ($p > 0.05$) coefficients are marked with *.

ternational results of the data analysis show that if internal locus of control increases by one unit, the students' mathematics score increases by 16.6 score points (if external locus of control is constant). If external locus of control increases by one unit, then the students' mathematics score falls for 32.8 score points. Every unit increase in the external locus of control is therefore associated with 32.8 score points fall in the students' mathematics achievement (if the effect of internal locus of control is held constant). On an international level, the model accounts for 11 % of variance in the students' mathematics achievement score.

Likewise, the results of the data analysis for Slovenia show that if internal locus of control increases by one unit, the students' mathematics score increases by 15.3 score points. Therefore, every unit increase in the internal locus of control is associated with 15.3 score points increase in the students' mathematics achievement (if external locus of control is constant). If external locus of control increases by one unit, then the students' mathematics score falls for 28.3 score points. Every unit increase in the external locus of control is therefore associated with 28.3 score points fall in the students' mathematics achievement (if internal locus of control is constant). In Slovenia, the model accounts for 8 % of variance of the students' mathematics achievement score.

Further comparisons of the countries with the highest and lowest mathematics achievement scores in European Union showed that the regression model, which accounts for the highest percentage of variance (15 %), is the regression model for Estonia. The results of the data analysis for Estonia show that every unit increase in the internal locus of control is associated with 18 score points increase in the students' mathematics achievement (if external locus of control is constant). Every unit increase in the external locus of control is therefore associated with 36 score points fall in the students' mathematics achievement (if internal locus of control is constant). Moreover, the regression model for Romania accounts for the lowest percentage of variance (9%) in analysis. The results for Romania show that every unit increase in the internal locus of control is associated with 10 score points increase in the students' mathematics achievement (if the external locus of control is constant). Moreover, every unit increase in the external locus of control is therefore associated with 29 score points fall in the students' mathematics achievement (if the internal locus of control is constant).

Table 5 shows that the inclusion of an additional index of attribution for failing mathematics does not add to percentage of explained variance to the original regression model which includes only an attribution for success in mathematics indices. The inclusion of an additional index of

Table 5: Regression model with attribution for (perceived self-responsibility for failing mathematics - FAILMAT) index included

	b (s.e.)	β (s.e.)	R^2 (s.e.)
International results			
constant	477.65* (0.42)		
Internal locus of control	-16.01* (0.30)	-0.15* (0.00)	
External locus of control	30.84* (0.39)	0.25* (0.00)	
FAILMAT	-3.52* (0.29)	-0.04* (0.00)	0.11 (0.00)
Slovenia			
constant	504.75* (1.69)		
Internal locus of control	-15.38* (2.12)	-0.14* (0.02)	
External locus of control	26.56* (3.11)	0.21* (0.02)	
FAILMAT	-3.75 (2.12)	-0.04 (0.02)	0.08 (0.01)
Netherlands			
constant	529.82* (3.77)		
Internal locus of control	-15.13* (2.19)	-0.17* (0.02)	
External locus of control	56.94* (4.58)	0.26* (0.03)	
FAILMAT	1.28 (0.01)	0.01 (0.03)	0.11 (0.01)
Estonia			
constant	516.52* (2.33)		
Internal locus of control	-17.94* (2.02)	-0.18* (0.02)	
External locus of control	37.44* (2.57)	0.31* (0.02)	
FAILMAT	-1.04 (0.01)	-0.01 (0.02)	0.15 (0.01)
Romania			
constant	455.15* (3.86)		
Internal locus of control	-11.52* (2.37)	-0.11* (0.02)	
External locus of control	27.97* (3.20)	0.27* (0.03)	
FAILMAT	-2.66 (1.96)	-0.03 (0.02)	0.09 (0.02)

	b (s. e.)	β (s. e.)	R^2 (s. e.)
Bulgaria			
constant	444.48* (3.26)		
Internal locus of control	-12.72* (2.15)	-0.12* (0.02)	
External locus of control	41.45* (2.79)	0.37* (0.02)	
FAILMAT	0.17 (1.99)	0.00 (0.02)	0.15 (0.02)

Notes: The data are weighted with Final Student Weight. FAILMAT is an abbreviation for the index "perceived self-responsibility for failing mathematics". R^2 is adjusted R^2. All the data presented in this table are calculated using only the data for students who answered question ST43 (attribution for success in mathematics) in Student Questionnaire. Statistically significant ($p > 0.05$) coefficients are marked with *.

attribution for failing mathematics accounts for an additional 1 % only for Bulgaria. For the international data and the rest of the countries (Slovenia, Netherlands, Estonia and Romania), the percentage of variance explained stays the same after the inclusion of additional predictor. Therefore, it can be concluded that the inclusion of the new predictor has not explained a large amount of the variation in students' mathematics achievement scores. The attribution for failure (perceived self-responsibility for failing mathematics) is a weaker predictor for the students' mathematics achievement score than the predictors of the attribution for success.[6] Moreover, the predictor attribution for failure (perceived self-responsibility for failing mathematics) is statistically significant in predicting students' mathematics achievement scores only on the international level.

Discussion

Internal locus of control as measured in PISA study is a significant predictor of higher mathematics achievement on international level and based on the samples included also regardless of average levels of mathematics achievement (Slovenia, Netherlands, Estonia and Romania). Likewise external locus of control significantly predicts lower mathematics achievement on an international level and in selected countries. The results showed predictive stability – in other words the predictors were significant in all analysed countries. In Slovenia, the students' attribution style explains 8 % of the total mathematics achievement score indicating the relevance of the analysed field.

6 The analysis of the data gave the same results when changing the order of predictors and including the predictor of "FAILMAT" as the first predictor in the regression analysis.

Based on our results, internal locus of control should be supported in educational setting. Inclusion of an additional index of the attribution for failure (perceived control responsibility about failing math) does not increase prediction value to a larger extent. This indicates that the attribution for success is something we should pay attention to in the educational setting. For instance, teachers could focus on communicating praises for success in a matter that promotes effort (internal, instable controllable attributions).

Hence, the main question for our discussion is how is attributional information developed in the course of one's development. Besides one's own experiences and social norm information, which is the strongest source, also feedback from teachers is relevant to motivation in school, especially because teachers are often unaware of the attributional information that they indirectly convey. For instance, laboratory – experimental studies showed that three types of behaviour that a teacher frequently poses could be problematic in communicating low ability (internal, stable, uncontrollable attribution) of student. These are (Graham and Williams, 2009): sympathy following failure, the offering of praise following success especially at easy tasks and unsolicited offers of help. In classroom teachers reward the effortful (internal, unstable, controllable attribution) student and punish the non effortful and unmotivated students. When a teacher attributes student's failure to lack of effort the student is perceived to be responsible, anger is elicited, and punishment or reprimand is handed out. In contrast, when failure is attributed to low aptitude and the student is perceived as not responsible sympathy is aroused, and help may be offered (Graham and Williams, 2009; Reyna and Weiner, 2001; Rudolph et al., 2004). That is, one tends to be sorry for the student who lacks ability, or is sick or breaks down on the way to school (Reyna and Weiner, 2001). In contrast to failure, being attributed to controllable causes such as lack of effort tends to evoke anger with withdrawal of help. This means that if a student experiences sympathy when faced with failure they also interprets this sympathy as attributional information stating that the event (e.g. failure) is uncontrollable. Unsolicited help has the same effect. This effect is evident even when students of different ages observe teacher behaviour toward other students. For instance, when observed on tape: the student that was given unsolicited help from their teacher was perceived as less able in comparison to their classmate that did not receive help from the teacher (Graham and Baker, 1990). Studies (among college students and children) showed that students who were praised for success at a relatively easy task were inferred to be of a lower ability in comparison to their classmates who were given neutral feedback (Graham and Williams, 2009).

We are not suggesting that these types of teacher's feedback always work but some critical attention has to be put also to this aspect especially in regards of unsolicited help. Stepping back and not providing help if not directly asked is supported based on the theoretical assumptions and also our data indicating that internal locus of control should be supported. Not providing help or waiting a little bit longer than usual not only supports controllable internal attribution styles but also supports autonomy as one of the basic foundations of inner motivation (Ryan and Deci, 2009).

Besides informing teachers in the form of teacher training on specific competences and on effort praise depending on the task difficulty, also student trainings have proven to be successful. In these types of intervention, teachers or other trained professionals guide students towards attributing failure to lack of effort. Dweck (1975 in Graham and Williams, 2009) has shown that students that have had helpless attributions (e.g. stable low ability) and have gone through attribution retraining have shown more persistence and more effort in future tasks compared to their control group (students of similar attribution style but without attribution retraining). Attribution based intervention have for instance in a group of college students resulted in 18 % higher rate in passing the final exam (Van Overwalle and De Metsenare, 1990). Nevertheless, all the studies have not yielded such promising results (Stroud and Reynolds, 2009) and additional research in the field is needed.

It is also recommended that training should be subject-area specific – as our empirical study was targeting only mathematics achievement. Vispoel and Austin (1995) showed a systematic trend for external attributions to generalize across subject areas and for internal attributions to remain subject-area specific. In school, elementary students' current and future attributions perceptions have been found to vary daily across assignments (Shell and Husman, 2008). It is of high importance to add that attribution beliefs are only one piece of puzzle in achievement motivation theoretically and empirically interrelated with other motivational construct such as goal orientation and affect. Shell and Husman (2008) pointed out that we cannot simply expect a rise in achievement solely by influencing one component of motivation.

Taking Ajzens' theory (OECD, 2013b) of planned behaviour as a framework (this framework was used in the development of PISA 2015 questionnaires), we can assume that by influencing internal control belief, we can foster one's behaviour e.g. mathematics related effort, mathematic related student behaviour and indirectly if possible students mathematics outcome. Even the theory of planned behaviour states that volitional be-

haviour is determined by specific attitudes and subjective norms together with perceived control (OECD, 2013b).

To sum up, our study firstly offers two newly developed indices based on the PISA 2012 question set measuring attribution for success in mathematics (*internal locus of control* and *external locus of control*) over one's success. These two indices could be used in further analyses in the field and also in the data sets not included in our analyses. The study offered empirical international support for the significant relationship between internal locus of control and higher academic achievement (in our case PISA mathematics achievement) and likewise external locus of control and low achievement.

Despite the contributions offered by this study, its limitations should also be noted. First of all, the measurement of the attribution for success is limited to six items therefore the findings should be considered as a form of screening and be used as a baseline for further more in depth measurement of attribution style. The study included only a selection of countries therefore the conclusions could be generalized to other European countries to a certain extent but keeping in mind that the results are based on five selected countries. Based on our results, we can recognize the possibility of larger predictive value of attribution style in low achieving countries when compared to high achieving countries. However, in order to make these kind of conclusions we would have to include the whole PISA sample and conduct more in depth analyses. In the regression model, only newly developed indices were used (together with the attribution for failure in mathematics) mainly because we wanted to isolate the predictive power of the attributions for success in mathematics for mathematics achievement but at the same time this means that larger regression models could explain achievement in larger extent.

To conclude, success in school depends on numerous factors, many of which are not fully controllable or easily identified. Therefore, it is of vital importance that we identify factors that we can influence and, in such a way, help the students to reach better educational achievements on all levels and in all fields.

References

American Institutes for Research (2008) Plausible Value Procedures. Available from: http://www.rasch.org/rmt/rmt182c.htm (Accessed 20th October 2014).

Bong, M. (2004) Academic motivation in self-efficacy, task value, achievement goal orientations, and attributional beliefs. *Journal of Educational Research.* 97, pp. 287-297.

Carr, M., Borkowki, J. G., and Maxwell, S. E. (1991) Motivational components of underachievement. *Developmental Psychology.* 27, pp. 108-118.

Georgiou, S. N. (1999) Achievement attributions of sixth-grade children and their parents. *Educational Psychology.* 19 (4), pp. 399-412.

Gibb, B.E., Zhu, L., Alloy, L.B., and Abramson, L. Y. (2002) Attributional styles and academic achievement in university students: a longitudinal investigation. *Cognitive therapy and research.* 26 (3), pp. 309-315.

Graham, S., and Barker, G. (1990) The downside of help: An attributional–developmental analysis of help giving as a low ability cue. *Journal of Educational Psychology.* 82, pp. 7-14.

Graham, S., and Williams, C. (2009) An attributional approach to motivation in school. In: Wentzel, K.R., and Wigfiled, A. (eds.). *Handbook of motivation in school.* New York and London: Routledge, pp. 11-33.

Field, A. (2000) *Discovering statistics using SPSS for Windows.* London: Sage publications.

Khodayarifard, M., Brinthaupt, T. M., and Anshel, M. H. (2010) Relationships of parents' and child's general attributional styles to academic performance. *Social Psychology Education.* 13, pp. 351-365.

Liu, K. S., Cheng, Y. Y., Chen, Y. L., and Wu, Y. Y. (2009) Longitudinal effects of educational expectations and achievement attributions on adolescents' academic achievements. *Adolescence.* 44 (176), pp. 909-924.

Mattews, G., Deary, I. J., and Whiteman, M. C. (2009) *Personality traits.* Cambridge, England: Cambridge University Press.

Mcclure, J., Meyer, L. H., Garisch, J., Fischer, R., Weir, K. F., and Walkey, F. H. (2011) Students' attributions for their best and worst marks: Do they relate to achievement? *Contemporary Educational Psychology.* 36, pp. 71-81.

NCES (2014a) Program for International Student Assessment. Available from:http://nces.ed.gov/surveys/pisa/faq.asp (Accessed 20th October 2014).

NCES (2014b) Methodology and Technical Notes – Scaling of Data. Available from:http://nces.ed.gov/surveys/pisa/pisa2012/pisa2012highlights_9e.asp (Accessed 20th October 2014).

Nokelainen, P., Tirri, K., and Marenti-Valimaki, H. (2007) Investigating the influence of attribution styles on the development of mathematical talent. *Gifted child quarterly.* 51 (1), pp. 64-80.

OECD (2013a) *PISA 2012 Results: Ready to Learn: Students' engagement, Drive and Self-Beliefs (Volume III).* Paris: OECD Publishing.

OECD (2013b) *PISA 2012 Assessment and Analytical Framework.* Paris: OECD Publishing.

OECD (2013c) *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy.* Paris: OECD Publishing.

OECD (2014) *PISA 2012 Results: What Students Know and Ca Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014).* Paris: OECD Publishing.

Reyna, C., and Weinar, B. (2001) Justice and utility in the classroom: an attributional analyses of the goals of teachers punishment and intervention strategies. *Journal of educational psychology.* 93 (2), pp. 309-319.

Rudolph, U., Resch, S.C., Greitemeyer, T., and Weiner, B. (2004) A meta–analytic review of help giving and aggression from an attribution perspective. *Cognition and emotion.* 18, pp. 815-848.

Ryan, R.M., and Deci, E.L. (2009) Promoting self-determined school engagement: motivation, learning and well-being. In: Wentzel, K. R. and Wigfiled, A. (eds). *Handbook of motivation at school.* New York: Routledge, pp. 171-197.

Sarafino, E.P. (2006) *Health Psychology: bio psychosocial interactions (5th edition),* New York: Wiley.

Schunk, D. H. (1982) Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology.* 74 (4), pp. 548-556.

Shell, D.F., and HUSMAN, J. (2008) Control, motivation, affect, and strategic self-regulation in the college classroom: a multidimensional phenomenon. *Journal of educational psychology.* 100 (2), pp. 443-459.

Stipek, D. J., and Decotis, K. M. (1988) Children's understanding of the implications of causal attributions for emotional experiences. *Child Development.* 59, pp. 1601-1610.

Stipek, D. J., and Hoffman, J. M. (1980) Children's achievement-related expectancies as a function of academic performance histories and sex. *Journal of Educational Psychology.* 72, pp. 861-865.

Stroud, K., and Reynolds, C.R. (2009) Assessment of learning strategies and related constructs in children and adolescents. In: Gutkin, T.B., and Reynolds, C.R. (eds.). *The handbook of school psychology.* Hoboken: Wiley and sons, pp. 739-765.

Van Overwalle, F., and De Metsenaere, M. (1990) The effects of attribution-based intervention and study strategy training on academic achievement in college freshmen. *British Journal of Educational Psychology.* 60, pp. 299-311.

Vispoel, W. P., and Austin, J. R. (1995) Success and failure in junior high school: A critical incident approach to understanding students' attributional beliefs. *American Educational Research Journal.* 32, pp. 377-412.

Weiner, B. (1985) An attributional theory of achievement motivation and emotion. *Psychological Review.* 92, pp. 548-573.

Weiner, B. (2010) The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist.* 45, pp. 28-36.

Zimmerman, B. J. (2000) Attaining self-regulation: A social cognitive perspective. In: Boekaerts, M., Pintrich, P. R., and Zeidner, M. (eds.). *Handbook of self-regulation.* San Diego, CA: Academic Press, pp. 13-39.

(In)equalities in PISA 2012 mathematics achievement, socio-economic gradient and mathematics-related attitudes of students in Slovenia, Canada, Germany and the United States

Mojca Štraus

Introduction

The associations of student's socio-economic background with their educational achievement have long been well established (e.g. Duru-Bellat, 2004). For decades, international large-scale assessments such as the Programme for International Student Assessment (PISA) discovered significant relationships between the family background and student's performance. For example, the international PISA report on equity in schools describes that "socio-economically advantaged students and schools tend to outscore their disadvantaged peers by larger margins than between any other two groups of students" (OECD 2013a, p. 34). However, PISA also showed that higher achievement is not necessarily at the expense of equality: "Many countries and economies that have seen improvements in their mean performance on PISA have also managed to weaken the link between the socio-economic status and performance, sometimes resulting from a narrowing of the gap in performance between advantaged and disadvantaged students" (ibid, p.35).

There is evidence of the impact of socio-economic background on mathematics achievement in Slovene schools as well. In addition to the results that can be observed from large-scale international assessment, for Slovenia, Žakelj and Grmek (2010) demonstrated associations between students' socio-economic status and their achievement in the National Assessment of Knowledge (National Examinations Centre 2014), conducted every year at the end of compulsory education. Also, Štraus and Markelj (2011) showed that the choices students make from academic, technical and vocational upper secondary educational programs tend to parallel social lines. This warrants further studies into the relationship between students' socio-economic background and achievement in Slovenia.

However, socio-economic status is not the only factor influencing achievement. How students feel and think about themselves is an important predictor of how they act and decide when challenged by tasks and situations (Bandura 1977). There is vast research on students' attitudinal factors influencing mathematics achievement (e.g. Lamb and Fullarton 2001, Ma and Bradley 2004, Pang and Rogers 2013, Suan 2014). These factors can originate in students' home or school environment, or stem from their own perceptions.

The aim of this article is to examine to what extent students' attitudinal factors additionally explain the relationship between socio-economic background and achievement. In the study, mathematics achievement is taken into consideration, as it was the major domain in the latest cycle of PISA, in 2012, and it is one of the core subjects assessed in the National Assessment of Knowledge every year. In order to investigate the relationship between family background and student achievement, Willms (2003, 2006) introduced a concept of socio-economic gradient. PISA uses a composite measure of students' economic, social and cultural background derived from the highest occupational status and educational level of parents and home wealth, cultural and educational possessions (e.g. OECD 2013b). The socio-economic gradient is used to describe the quality of educational outcomes as well as the equality of outcomes for students from differing socio-economic backgrounds.

During PISA assessment, students are given a questionnaire including questions about their perceptions that are conceptualized in PISA to be relevant in explaining differences in their achievement (OECD 2012). Some of these data describe students' current school environment, for example sense of belonging to school or students' views of teacher behaviour. In Slovenia, this is a school environment in which students have generally only been part of for a few months – over 90% of 15-year-olds in Slovenia are in Year 1 of upper secondary school (Statistical Office of the Republic of, 2014a, 2014b). As such, these data are not likely to describe the context of the students' learning development in the previous years. Rather, inclusion of the indicators on current school environment in the model of factors relating to mathematics achievement in Slovenia could possibly mask other, more subtle factors from previous development. As described, the choices students make for their upper secondary programs parallel their social status and their responses to questions about their current school environment may, to a large extent, reflect this phenomenon.

The focus of analysis in this article is therefore limited to the constructs that more likely pertain to the students' background during a longer period. Such constructs are, for instance, self-concept in mathe-

matics, anxiety about mathematics or perseverance. Such student-level constructs from the attitudinal domain generally play a two-fold role in educational assessments. First, they are seen as the pre-conditions for success in subject domains, and, second, they themselves may be judged as educational goals (Rychen and Salgalnik, 2003). The constructs of interest in this study are taken from Ajzen's (1991, taken from OECD 2012, p.185) theory of planned behaviour which states that by manipulating attitudes and subjective norms as pre-determinants of volitional behaviour, the chances that the person will intend to do a desired action can be increased which, in turn, can increase the likelihood of the behaviour actually occurring. By introducing attitudinal constructs in the model of relationship of socio-economic background and mathematics achievement, we try to investigate further the interplay of these constructs with achievement.

To provide a perspective or a reference point for the results of proposed investigations, other countries are included in the analysis. It is known from PISA (and other international studies) that there is a strong relationship between the students' family background and their achievement in all participating countries, however, there are possible differences in the factors that play a role in explaining this relationship in one country and not another. A comparative perspective therefore seems useful. The countries selected for comparisons are Canada, Germany and the United States. While it is obvious that the choice of countries is based on the content of the present thematic issue having guest authors from these three countries, it also gives a good framework for interpreting the results for Slovenia. From the educational policy perspective, international comparisons of Slovene students' achievements are an important indicator of educational quality. According to PISA 2012 international reports, mathematics achievement in Canada and Germany is higher than in Slovenia and in the United States it is lower (OECD, 2013c). At the same time socio-economic status of students varies among the countries; it is highest in Canada, similar between Germany and the United States and lowest in Slovenia[1]. Contextually, all three countries are major world economies and Germany is the strongest economic partner of Slovenia (Observatory of Economic Complexity, 2014). So these countries present different viewpoints from which results for Slovenia can be interpreted.

From the above discussion, the following research question is formulated: *To what extent does variation in mathematics-related attitudes of students, in addition to their socio-economic background, contribute to explaining their mathematics achievement and how do these relationships compare between Slovenia, Canada, Germany and the United States?*

1 These data are presented in more detail later in this article.

This research question proposes studying the distribution of mathematics achievement by student-level socio-economic and attitudinal factors in Slovenia and three other countries. It needs to be noted that the PISA assessment measures a broad set of skills and knowledge in mathematics, called mathematical literacy of 15-years-old students enrolled in formal education (OECD, 2012). Therefore, it is not an assessment of what students learned in mathematics classes during their previous year, or during their lower secondary education. It is an assessment of the cumulative development of learning that has occurred throughout education or elsewhere.

The next section describes constructs, data and methods used in this article. Then, results of comparisons of socio-economic gradients in mathematics achievement in Slovenia with the other three countries will be given, followed by the results of analyses of socio-economic background's interrelationships with students' attitudes in describing mathematics achievement of 15-years-old students. These results are discussed in the final section.

Data and Methods

The latest PISA data collection was conducted in 2012 with mathematics literacy as the major domain. The majority of items in achievement tests therefore covered mathematics and the questionnaire items asked about constructs conceptualized in PISA to be related to mathematics achievement. These data will be used in the present study.

Samples of Students

The PISA data collection included large nationally representative samples of 15-years-old students, consistent with the requirements for the international study (OECD, 2013c1). Furthermore, PISA 2012 introduced a rotation design for the student questionnaire, similar to the design for cognitive items. This means that questionnaire items, organized in item packages, were distributed over three forms of questionnaire booklets, named Form A, Form B and Form C. Each student was allocated one of these booklets, and thus received a limited number of questionnaire items, while the booklets taken together covered a larger, "universal" pool of questionnaire items. Like the cognitive booklets, the questionnaire booklets were randomly rotated among students within each school to make sure representative samples of students took each of the questionnaire forms (OECD forthcoming). Each form was therefore allocated to approximately a third of students in the PISA sample in a particular country.

Table 1: Differences in sample sizes and population estimates based on full PISA 2012 samples and Form B Questionnaire subsamples[2]

	Slovenia	Germany	Canada	United States
Number of students in full PISA 2012 sample	5911	5001	21544	4978
Number of students in Form B subsample	1896	1654	7173	1665
Mean achievement based on full PISA sample	501 (1.3)	514 (2.9)	518 (1.8)	481 (3.6)
Mean achievement based on Form B subsample	505 (2.6)	518 (3.2)	519 (2.4)	484 (4.1)

Due to this rotation, not all students provided responses for all of the attitudinal factors to be used in this study. Data on the full set of items of interest was collected through Form B of PISA 2012 student questionnaire only. In order to analyse the factors derived from these items within a single model, we therefore used data from subsamples of 15-years-old students that were given Form B during the PISA assessment from each country. Consequently, population estimates based on these subsamples may slightly differ from the estimates obtained from full PISA samples that are available in PISA international reports. The two sets of estimates are presented in Table 1 for comparison. The differences between the full sample and the Form B-subsample estimates for Germany and Slovenia are somewhat larger due to the fact that in these two countries, small subgroups of students were given a shorter version of the cognitive assessment as well as the questionnaire, named UH (une heure) instruments. Such versions are available in PISA for students with special educational needs that otherwise could not participate in the PISA assessment. This version of instruments was given to 3% of students in Germany and 1% of students in Slovenia. Exclusion of these students from the calculations of the mean achievement in a country generally results in an increased estimate. However, UH instruments did not collect data on any of the factors selected for the present study, except for the index of socio-economic and cultural status, so students that were given these instruments would not be included in the analysis in any case. As none of the four differences between full-sample mean estimates and the Form B subsample estimates

2 In the whole article, standard errors are given in parentheses. The standard errors indicate the accuracy of the estimates. For example, if one imagines that the PISA study had been repeated a number of times with the same sample sizes for each country, then in about 95% of cases, the estimates of the means would have fallen within the double range indicated by the standard errors

are significant[5], the data from Form B subsamples were deemed of sufficient quality to be used for the present investigation. From here on, it is only these data that are included in the analysis.

Mathematics-related Constructs

Table 2: Structure of blocks and indices of students' mathematics-related attitudes and opinions

Block 1 Socio-economic background	
Index of socio-economic and cultural status (ESCS)	Constructed index based on students' responses about their parents' education and occupation and home possessions
Block 2 Mathematics self-beliefs and participation in mathematics-related activities	
Index of mathematics self-efficacy	Constructed index based on students' responses about their perceived ability to solve a range of pure and applied mathematics problems
Index of mathematics self-concept	Constructed index based on students' responses about their perceived competence in mathematics
Index of mathematics anxiety	Constructed index based on students' responses about feelings of stress and helplessness when dealing with mathematics
Index of subjective norms in mathematics	Constructed index based on students' responses about whether they intend to use mathematics in their future and whether students' parents and peers enjoy and value mathematics
Block 3 Students' drive and motivation	
Index of perseverance	Constructed index based on students' responses about their willingness to work on problems that are difficult, even when they encounter problems
Index of openness to problem solving	Constructed index based on students' responses about their willingness to engage with problems
Index of perceived self-responsibility for failing in mathematics	Constructed index based on students' responses about whether they attribute failure in mathematics tests to themselves or to others
Index of intrinsic motivation to learn mathematics	Constructed index based on students' responses about whether they enjoy mathematics and work hard in mathematics because they enjoy the subject
Index of instrumental motivation to learn science	Constructed index based on students' responses about whether they believe mathematics is important for their future studies and careers

In PISA 2012, background data was collected with the aim to portray important aspects of the affective domain, such as valuing mathematics and being confident in doing mathematics. From the data that were collected via student questionnaires, interval-scaled statistical indices were derived to capture the major constructs related to mathematics achievement.

5 Tested following the procedures in OECD (2009).

There is a set of indices given in the PISA 2012 database from which it is possible to select the indices that basically capture the major aspects in Ajzen's (1991) theory of planned behaviour. The indices selected for the present study are organized in three blocks; the first block comprises of a single index of socio-economic and cultural status, the second block comprises of indices of students' mathematics-related self-beliefs and the third block of indices of their drive and motivation in mathematics. Descriptions of these indices are presented in Table 2. Concrete items in the PISA 2012 student questionnaires that were used to collect data for the selected indices and data for these items are detailed in OECD (2013b).

A statistical index in the PISA database is constructed in a way that for all students in the OECD countries the mean is 0 and the standard deviation 1 (in computing the mean and standard deviation an equal weight is given to each of the participating countries) (OECD, 2013b and OECD forthcoming). Negative values of the index in the international database therefore do not imply that students responded negatively to the underlying questions, but rather that they responded less positively (or more negatively) than the average response across OECD countries. Likewise, positive values imply more positive (or less negative) responses than the average response in OECD countries.

Socio-economic Gradient

Willms (2003) describes that socio-economic gradients comprise of three components, mean level, mean slope and the strength of the relationship between the outcome variable and socio-economic background. The *level* of the gradient is defined as the expected score on the outcome measure for a student with average socio-economic status. The level of a gradient for a country is an indication of its overall performance, after taking into account the students' socio-economic status. The *slope* of the gradient is an indication of the extent of inequality attributable to socio-economic status. Steeper gradients indicate a greater impact of socio-economic status on student performance (greater inequality) while gradual gradients indicate lower impact of socio-economic status (less inequality). The *strength* of the gradient refers to how much individual scores vary above and below the gradient line. If the relationship is strong, then a considerable amount of the variation in the outcome measure is associated with socio-economic status, whereas a weak relationship indicates that relatively little of the variation is associated with socio-economic status. The most common measure of the strength of the relationship is a statistic called R-squared, which is the proportion of variance in the outcome measure explained by the predictor variable.

Statistical Analyses

The main analytical approach for the investigation in this article is linear regression analysis, conducted in a sequence of steps. First we estimate socio-economic gradient using a simple one-predictor model for each of the four countries. Then the model is extended with factors capturing various aspects of students' mathematics-related attitudes. The appropriate structure of these factors for the final model is derived from preliminary exploratory analyses. For all four countries the same final model is used.

Due to the clustering structure of the PISA data – students being sampled within previously sampled schools – the question whether hierarchical modelling needs to be used should be addressed. Since only student-level variables are investigated in our study separately for each of the four countries, it remains to be considered whether the variance of these variables shared between the schools is of interest. The impact of clustering on sampling variance is controlled for by Bootstrap procedures of computation.

As mentioned, the majority of 15-year-old students in Slovenia attend the first year of their upper secondary education segregated to different educational programs and that the students' selection of these programs tends to parallel their socio-economic background. Therefore it seems self-evident that the proportion of variance in mathematics achievement as well as other variables between schools is relatively large. The linear regression coefficient of socio-economic background on the student achievement provides an estimate of the overall difference in performance due to socio-economic background while multilevel regression model estimates the difference in performance after accounting for the differential attendance to schools. The multilevel regression coefficients on socio-economic background may therefore substantially differ from the linear regression coefficients, especially in highly tracked systems. Having four different education systems in our study, the primary interest are the overall differences in the populations of students while differences between schools are left aside. For this reason, the linear regression is used.

IBM SPSS 22.0 software is used for the analyses, with the addition of the syntax macros prepared through the IDB Analyzer software (IEA, 2014), which enables calculations of population estimates and standard errors with the use of suitable sample weights and all five plausible values of achievement in the PISA database. Throughout the article, significance of differences in mean estimates or in estimates of regression coefficients between countries is tested using the foundations in OECD (2009). Testing is carried out at 0.05 level of statistical significance between results for Slovenia and each of other countries.

A final note of caution is in order. When interpreting the results of investigations in this article, it should be taken into consideration that the indices used in the analyses have been derived from students' responses to questions in the background questionnaire and not from, for example, independent observations or other types of objective measurements. This means that students' answers depended on the way students understood and responded to questions.

Results

International Comparisons of Slovene Students' Mathematics Achievement and its Socio-economic Gradient

From PISA 2012 data, we can derive basic comparisons of mathematics achievement and its socio-economic gradient between Slovenia, Canada, Germany and the United States. While these indicators are available in the PISA initial reports (e.g. OECD, 2013a), it is important to repeat that this study uses subsamples of the original PISA samples within the selected countries and, consequently, some of the indicators in Table 3 slightly differ from the initial reports.

Table 3: Data on socio-economic gradient in mathematical literacy for Slovenia, Germany, Canada and the United States in PISA 2012

	Slovenia	Germany	Canada	United States
Mean socio-economic and cultural status	0.07 (0.03)	0.21 (0.03)	0.41 (0.02)	0.16 (0.04)
Mean score in mathematical literacy	505 (2.6)	518 (3.2)	519 (2.4)	484 (4.0)
Level of socio-economic gradient1	521 (2.4)	521 (3.6)	515 (2.4)	489 (3.3)
Slope of socio-economic gradient2	45 (2.6)	42 (3.5)	34 (2.1)	36 (2.6)
Strength of socio-economic gradient3	16.1 (1.7)	13.8 (1.9)	10.2 (1.1)	14.3 (1.9)

Notes: 1 Level of socioeconomic gradient is the mean score in mathematical literacy adjusted for the mean socio-economic and cultural status (ESCS). Adjusting for socio-economic and cultural status takes into account only mean achievement of groups of students with socio-economic and cultural status equal to OECD average in each country.
2 Slope of socio-economic gradient is the score-point change in achievement associated with one-unit increase in socio-economic and cultural status.
3 Strength of socio-economic gradient is the strength of the relationship between mathematical literacy and socio-economic and cultural status (ESCS) as the percentage of variance in mathematics performance explained by the socio-economic and cultural status.

First, Table 3 shows differences between the four countries in average socio-economic and cultural status. Average socio-economic and cultural status of Slovene 15-years-old students (value 0.07) is slightly above

the OECD average (which is 0[4]), however, it is the lowest value of the four countries. Socio-economic and cultural status of 15-years-old students in Canada is the highest (value 0.41) and the values for Germany and the United States are in-between (values 0.21 and 0.16, respectively).

The values of mean scores in mathematical literacy show a different pattern. The scores for Germany and Canada are similar, the score for the United States is the lowest and the score for Slovenia is in-between. This shows that the socio-economic and cultural status itself does not determine the level of mathematics achievement in a particular country. Furthermore, while one could try to argue that the level of achievement in Slovenia is understandably lower than in Canada and Germany due to lower socio-economic and cultural status of Slovenian students, this is not supported by the level of socio-economic gradient in Table 3. It can be observed that differences between Slovenia and the other countries still exist even when mathematics achievement is adjusted for students' socio-economic and cultural status. These comparisons show that the levels of socio-economic gradients indeed vary between the four countries.

Another element of the socio-economic gradient, the slope, also varies between the countries. One can observe that the slopes of socio-economic gradients in Slovenia and Germany are the two highest (45 and 42 points, respectively) and in Canada and the United States the two lowest (34 and 36 points, respectively). In other words, in Slovenia and Germany a one-unit increase in socio-economic and cultural status is associated with a somewhat higher increase in mathematics achievement than in Canada and the United States. It is interesting that, even though both, average socio-economic status as well as average mathematics achievement of Canadian students are different from these characteristics of the United States' students, the slopes of the socio-economic gradients are similar between the two countries. The same can be observed for Slovenia and Germany. The percentage of variance in mathematics achievement explained by socio-economic and cultural status is the lowest in Canada indicating the weakest gradient among the four countries. In Slovenia, the socio-economic gradient seems to be the strongest. However, given the relatively small percentages of variance in mathematics achievement explained by socio-economic and cultural status in all four countries, it seems reasonable to expect that there are other factors accounting for the variance in mathematics achievement of students in the selected countries. Some of these factors are investigated in the next section.

4 For all indices, OECD average is 0. See section on data and methods.

Socio-economic Gradient Together with Self-related Beliefs in Mathematics

In this section, we present the structure of (some of) the underlying factors associated with mathematics achievement in Slovenia in comparison with Canada, Germany and the United States. The model for socio-economic gradient is extended with mathematics-related attitudinal factors. First, descriptive data on these factors are presented in Table 4. As explained, all factors are derived on an interval scale with a mean value 0 for OECD countries and standard deviation of 1. Values presented in Table 4 are therefore readily comparable.

Table 4: Mean values of factors[5]

	Slovenia	Germany	Canada	United States
Mathematics-related self-beliefs				
Index of mathematics self-efficacy	0,33 (0,03)	0,33 (0,03)	0,13 (0,02)	0,17 (0,03)
Index of mathematics self-concept	-0,02 (0,03)	0,10 (0,03)	0,20 (0,02)	0,31 (0,03)
Index of subjective norms in mathematics	-0,23 (0,03)	-0,12 (0,03)	0,36 (0,02)	0,28 (0,03)
Index of mathematics anxiety	0,13 (0,02)	-0,21 (0,03)	0,06 (0,02)	-0,04 (0,03)
Students' drive and motivation				
Index of perseverance	0,07 (0,03)	-0,02 (0,03)	0,19 (0,02)	0,38 (0,03)
Index of openness to problem solving	0,11 (0,03)	0,18 (0,03)	0,14 (0,02)	0,24 (0,04)
Index of self-responsibility for failing in mathematics	0,21 (0,03)	0,16 (0,03)	-0,22 (0,02)	-0,40 (0,03)
Index of intrinsic motivation to learn mathematics	-0,21 (0,03)	-0,18 (0,03)	0,01 (0,02)	0,12 (0,03)
Index of instrumental motivation to learn science	-0,31 (0,03)	-0,24 (0,03)	0,14 (0,02)	0,08 (0,02)

The average values of indices of students' attitudes and opinions about mathematics vary between the four countries. The levels of self-efficacy in mathematics show an interesting distinction between the countries. In Slovenia and Germany, students express high levels of conviction about their capability to cope with certain mathematics tasks (values of 0,33 in both countries), while students in Canada and the United States seem to be less convinced in their capabilities (values 0.13 and 0.17, respectively). This finding seems to be in contrast with the result that Canadian

5 Testing of statistical significance of differences in this article is carried out between results for Slovenia and each of other countries. For brevity, interpretations of comparisons between other countries are made more generally without testing for significance. This testing can be carried out using standard errors provided with each estimate. Due to estimates being based on Form B subsamples only, standard errors are somewhat larger and less significant differences can be established. Interpretations are made as indications of results for which further, more detailed investigations seem warranted.

students achieve the highest scores in mathematics among the four countries. However, perhaps the more important question is how does self-efficacy relate to achievement within individual countries. This will be presented later in this section.

A somewhat broader sense of the overall perception of students' personal attributes in mathematics, the mathematics-related self-concept, also varies between the countries, although rankings changed. Slovene students report the lowest self-concept, around the OECD average (value -0.02). German students report slightly higher values than the OECD average (value 0.10), with Canadian students reporting also higher (value 0.20), and the highest reports coming from students in the United States (value 0.31). Given that achievement in the United States is the lowest among the four countries, it is difficult to imagine this factor to be positively related to student achievement. However, it needs to be kept in mind that these are average values per country and that there is variation within individual countries in achievement as well as in the background factors. Again, this is examined later in this article.

A few additional indices portray clustering of values for Slovenia and Germany together on one side and of values for Canada and the United States on the other. The index of subjective norms captures the beliefs of student that specific individuals or groups think they should perform well in mathematics and students' motivation to comply with these groups. German and Slovene students expressed lower than average levels of such beliefs (values -0.12 and -0.23, respectively), and students from Canada and the United States well above average beliefs (values 0.36 and 0.28, respectively). The valuing of mathematics in the students' environment, as measured through the index of subjective norms, as well as intrinsic and instrumental motivation to learn mathematics are therefore relatively low in Slovenia and Germany and relatively high in Canada and the United States.

The index of self-responsibility for failing in mathematics reflects students' perceptions of their personal responsibility for failure in mathematics. Students with high values on this index tend to attribute the responsibility for failure to solve mathematics problems to themselves while students with low values on this index are more likely to see other individuals or factors as responsible. While students in Slovenia and Germany report relatively high levels of self-responsibility for failing in mathematics, students' reports show lower levels of this responsibility in Canada and the United States. Similarly, students in Slovenia and Germany report around average levels of perseverance, but students in Canada and the United States report higher perseverance. It is only for the openness for

problem solving that students in all four counties give closer reports, all of them above average. The index that most stands out from this pattern is mathematics anxiety. While students in Canada and the United States report around average levels of anxiety (values 0.06 and -0.04, respectively), German students report relatively low anxiety (value -0.20) but Slovene students report the highest levels of mathematics anxiety among the four countries (value 0.13).

In summary, among the four countries, Slovene students express the lowest self-concept in mathematics, the lowest intrinsic as well as instrumental motivation to learn mathematics and the lowest level of beliefs that their parents and peers think they should perform well in mathematics. At the same time they express the highest level of self-responsibility for failing in mathematics and the highest mathematics anxiety. This in itself is an important message about Slovene mathematics education.

Preliminary Regression Analysis

In the preliminary analysis, it was first explored which factors, individually or as blocks, explain most variance in mathematics achievement.[6] The results of this analysis showed that for all countries, self-efficacy explains more variance in mathematics achievement than socio-economic and cultural status[7]. When blocks of indices were entered into the model separately, it was found that a larger amount of variance is explained by the block of mathematics-related self-beliefs (between 28 and 36 percent for the four countries) than by the block of indices on drive and motivation (between 7 and 19 percent for the four countries). Furthermore, when both blocks were entered, the amount of variance explained was nearly the same as the amount of variance, explained only by the block of indices of mathematics-related self-beliefs[8]. From this, it was decided to use only the block of mathematics-related self-beliefs in the regression model.

In addition, issues of multicollinearity of the factors in the block of self-beliefs were tested. It was found that there are relatively large (negative) correlations between self-concept and anxiety (from -0.76 in the United States to -0.61 in Slovenia). Due to self-concept having larger (positive) correlations with self-efficacy than anxiety (from 0.39 in Slovenia to 0.55 in Canada), it was decided that anxiety is kept as the predictor

6 This was explored using a stepwise procedure for linear regression analysis in SPSS. Also other preliminary analyses were carried out in SPSS.

7 In a single-predictor model, socio-economic and cultural status explained between 15 and 17 percent of variance and mathematics self-efficacy explained between 25 and 35 percent of variance for the four countries.

8 The largest increase in amount of variance explained by adding both blocks of indices into the model was 2.5 percent.

while self-concept is dropped. Correlation analysis of the remaining factors showed that they correlate weakly or moderately ($-0.47 < r < 0.29$). Additionally, variance inflation factors (VIF) were significantly below 10 ($1.017 < VIF < 1.42$). Other research, however, shows that these concepts are different (Ferla et al. 2009) and have differential impacts on achievement across countries (Morony et al. 2012).

Results of Regression Analysis

A linear model[9] was set up in order to investigate differences in the impacts of selected factors on student mathematics achievement between the four countries. The results of regression analysis based on this model are presented in Table 5.

Table 5: Relationship between mathematical literacy, socio-economic and cultural status and mathematics-related self-beliefs[10]

Slovenia	b	¬	t(b)	t(□)	R^2	R^{2*}
constant	495 (2.3)		219.6			
ESCS	31 (2.8)	0.30 (0.02)	11.0	12.1		
MATHEFF	31 (2.8)	0.35 (0.03)	11.1	13.1		
SUBNORM	-8 (2.5)	-0.09 (0.03)	-3.3	-3.4		
ANXMAT	-19 (2.8)	-0.20 (0.03)	-6.8	-7.1	0.34	0.34
Germany	b	⊐	t(b)	t(□)	R^2	R^{2*}
constant	502 (2.4)		207.0			
ESCS	27 (2.1)	0.27 (0.02)	13.1	13.9		
MATHEFF	39 (2.9)	0.40 (0.03)	13.6	15.6		
SUBNORM	-14 (2.3)	-0.14 (0.02)	-6.0	-6.1		
ANXMAT	-16 (2.1)	-0.19 (0.03)	-7.3	-7.3	0.41	0.41
Canada	b	¬	t(b)	t(□)	R^2	R^{2*}
constant	512 (1.9)		276.0			
ESCS	21 (1.7)	0.21 (0.02)	12.8	12.6		
MATHEFF	37 (1.4)	0.43 (0.02)	26.1	28.8		
SUBNORM	-5 (1.4)	-0.06 (0.02)	-3.6	-3.6		
ANXMAT	-17 (1.5)	-0.19 (0.02)	-11.3	-12.1	0.38	0.38

9 The names used are the following ESCS – Index of socio-economic and cultural status; MATHEFF – Index of mathematics self-efficacy; SUBNORM – Index of subjective norms in mathematics; ANXMAT – Index of mathematics anxiety. The final model has the equation $MATH\_ACHIEVEMENT = a + b_1 \cdot ESCS + b_2 \cdot MATHEFF + b_3 \times SUBNORM + b_4 \times ANXMAT + error$

10 Coefficients presented in this table are all significant.

United States	b	⊓	t(b)	t(⊓)	R^2	R^{2*}
constant	479 (2.7)		178.5			
ESCS	23 (2.5)	0.26 (0.02)	11.5	11.3		
MATHEFF	35 (2.8)	0.42 (0.03)	12.7	13.7		
SUBNORM	-14 (1.9)	-0.16 (0.02)	-7.4	-7.4		
ANXMAT	-21 (2.5)	-0.24 (0.03)	-8.4	-8.4	0.41	0.41

With the model, it was possible to explain from 34 to 41 percent of variance in mathematics achievement in the four countries, seemingly the least in Slovenia[11]. In all countries, the mean achievements adjusted by the four predictors are closer together than the unadjusted means but the ranking of countries is the same. If four average students with regard to socio-economic and the selected attitudinal factors are taken from each of the countries, than the expected mathematics score is the highest for the Canadian student, 512 points, for the German student 502 points, for the Slovenian student 495 points and for the student from the United States 479 points. Also, by controlling the factors of the students' self-beliefs in the model, the socio-economic gradient becomes more gradual in all countries. For example, while the socio-economic gradient in mathematics achievement in Slovenia is 45 points (see Table 3), controlling for students' self-beliefs reduces the gradient to 31 points. This gradient indicates that if two groups of Slovene students with the same self-beliefs but one with a one-unit higher socio-economic and cultural status are compared than the higher-status group has on average 31 points higher mathematics achievement. Or in other words, even though students may have the same high or low mathematics-related self-beliefs, the ones with higher socio-economic and cultural status are, on average, expected to achieve higher in mathematics. The order of the reduced socio-economic gradients in the regression model remains the same as is the order of gradients obtained from the single-predictor model (see Table 3). The reduced gradients in Slovenia and Germany are the two steepest and in Canada and the United States the two most gradual of the four gradients.

However, analysis showed that believing in one's own capability of solving certain mathematics tasks remains a factor of relatively high impact even when other factors are controlled. In Germany, Canada and the United States students with similar socio-economic and cultural background and similar levels of subjective norms and mathematics anxiety, but with a one-unit difference in the levels of mathematics self-efficacy have on average over 35 points different scores on PISA mathematics test;

11 No significant differences were established between estimated proportion of variance for Slovenia and Canada.

students with higher self-efficacy having higher scores. In Slovenia, the impact of self-efficacy seems to be somewhat smaller; 31 score points.[12]

Generally, the index on subjective norms in mathematics, that is, the students' beliefs that their parents and peers value mathematics, was conceptualized to act as a positive predictor in the sense that students with higher values on this index achieve at higher levels (OECD, 2012). Results in the international PISA reports show that for the overall impact of subjective norms on mathematics achievement, this is true in Canada and the United States where a one-unit increase in this index is associated with an 8-point average increase in mathematics achievement in Canada and a 4-point increase in achievement in the United States. In Slovenia, there is no significant association between subjective norms and achievement but in Germany a one-unit increase in the index of subjective norms is associated with a 13-point *decrease* in mathematics achievement (OECD, 2013b). German students reporting more valuing of mathematics in their personal environment have on average lower achievement.

When this index is included in the model in the present investigation, its impact on mathematics achievement when other factors are controlled, becomes negative in all four countries.[13] If two groups of students in these countries are compared, having similar socio-economic and cultural status and expressing similar self-efficacy, and mathematics anxiety, than the group reporting higher values of subjective norms have on average lower achievement.

Mathematics anxiety presents no surprise as a predictor in the model. As shown by the results in international PISA reports, it has, in general, a negative impact on the achievement of at least a 27-point decrease per one-unit of this index in the four countries considered here (OECD, 2013b). 'This impact reduces substantially when other factors in the model are controlled. 'The decrease in achievement per one-unit increase in anxiety when controlling for other factors is between 16 and 21 scale points'[14].

Discussion and Conclusion

The goal of educational policy and reform in most countries is to raise levels of literacy skills, while reducing disparities among citizens from differing subgroups, like social classes and ethnic groups. In this article, we ad-

12 Significance of difference can be established between the results for Slovenia and Germany and the results for Slovenia and Canada.

13 The impacts of subjective norms and mathematics anxiety range from a 5 to 14 point decrease in achievement per one-unit increase of the factor. 'The significance of differences between Slovenia and any other individual country could not be established.

14 'The significance of differences between Slovenia and any other individual country could not be established.

dressed the issue of social gradient in student mathematics achievement and mathematics-related attitudes in Slovenia in comparison to three other countries, Canada, Germany and the United States. The availability and quality of PISA data provided an opportunity to gain further understanding on how differences in socio-economic and cultural background of students along with students' mathematics-related self-beliefs affect student achievement in mathematics.

The international PISA reports showed that, of the four countries, Canada and Germany are top-achieving countries, with mean mathematics achievements significantly above the mean in Slovenia, and the United States with mean mathematics achievement significantly below Slovenia. An overview of data on socio-economic and cultural status and mathematics-related attitudes showed Slovene students' socio-economic and cultural status are the lowest among the four countries and that most indices of mathematics-related attitudes of Slovene students are similar to Germany and opposite to Canada and the United States. Standing out from this pattern is the level of mathematics anxiety that, by students' reports, is the highest in Slovenia, average in Canada and the United States and below average in Germany. This can be taken as an indication of the area that needs further research in Slovenia. Such research may reveal the background of the results observed in this study.

To investigate how these aspects of student background and attitudes relate to student achievement, we set up a linear model, first only investigating socio-economic gradient and later expanding the model with attitudinal factors. Besides being assessed as outcomes of mathematics education, these constructs can also assist in explaining differences in performance on the PISA mathematics assessment. It was presumed that some of the variation in mathematics achievement observed by socio-economic background may overlap with variation in students' self-beliefs about mathematics.

There are several interesting findings from the analysis in this study. With regard to socio-economic gradient, this study, as many previous studies, found that there are inequalities in performance in all four countries associated with students' family background. The results also show that the extent of these inequalities varies between the countries. The most gradual socio-economic gradient among the four countries is found in Canada, then the United States, and then Germany and Slovenia.[15] In a similar order, the socio-economic gradient is the weakest – that is, socio-economic and cultural background of students explains the smallest

15 No statistical significance between the results for Germany and Slovenia could be established.

percentage of variance in mathematics achievement – in Canada (11 percent), then United States and Germany (14 percent) and is strongest in Slovenia (16 percent).[16] For Slovenia, this indicates the importance of research in the area of equity in education, such as the present study, to further illuminate the background for the observed results.

A preliminary regression analysis showed that in all four countries the indices of mathematics-related self-beliefs of students, as a block, are meaningfully stronger predictors of mathematics achievement than the block of indices on drive and motivation. This was seen by a larger amount of variance explained by the block of mathematics-related self-beliefs than by the block of indices on drive and motivation. Furthermore, the single factor explaining the largest amount of variance in mathematics achievement is self-efficacy. This is in line with findings from other studies (e.g. Ferla et al., 2009). As founded by the work of Bandura (1997), this indicates that conviction of one's own capability to perform is closely connected to achievement, in a circular manner where stronger conviction leads to better performance and better performance reinforces convictions. In reverse, if students are not convinced in their abilities to accomplish particular academic tasks, they have a higher probability of underperforming, even though they may have the ability. This is because they may not put in the self-control and motivation needed to perform the tasks. Zimmerman (2000) showed that self-efficacy is an important predictor of common motivational outcomes, such as students' activity choices, effort, persistence, and emotional reactions, but that it is sensitive to subtle changes in students' performance context.

Our analysis showed that if socio-economic and other attitudinal factors in the model are controlled, mathematics-related self-efficacy is still a strong and important predictor of mathematics achievement. The analysis further showed that among the four countries, this predictor seems to have the least impact in Slovenia. In efforts to avoid the vicious cycle of self-fulfilling prophecy for students with low self-efficacy, further investigations of this phenomenon seem warranted. A plausible hypothesis about the reasons behind this phenomenon may be that Slovene students with relatively high efficacy do not perform as well or students with low efficacy perform better. Given that average efficacy is as high in Slovenia as is in Germany, it seems the first is more likely than the latter. Although this finding needs to be cross-checked with additional information, like policy documents or data from additional countries, it seems a reasonable hypothesis that one of the reasons for this phenomenon may

16 The significance of differences could be established between the results for Slovenia and Canada.

come from teachers' practices in giving feedback to students (Zupanc and Bren, 2010).

A finding from preliminary analysis is also that, next to self-efficacy, the strongest factor explaining most of the remaining variance in mathematics achievement in all four countries is socio-economic and cultural status. When the attitudinal factors in the model are controlled, the socio-economic gradient varies from 21 to 31 scale points, the two largest being in Slovenia and Germany.[17]

A notable finding is also that while other factors preserved their conceptualized positive or negative nature of impact on mathematics achievement in the model, the index of subjective norms changed to a negative impact in all four countries. The initial PISA results published in the international reports already indicated that the nature of this factor's impact on mathematics achievement varies between the countries; in 18 countries it is positive, in 30 negative and in 17 countries its impact is neutral (OECD, 2013b). The finding that the impact of this factor in our model is negative in all four countries may be interpreted that the students agreeing with items 'most of my friends do well in mathematics', 'most of my friends work hard at mathematics', 'my friends enjoy taking mathematics tests', 'my parents believe it's important for me to study mathematics', 'my parents believe that mathematics is important for my career' and 'my parents like mathematics' actually responded about the pressure they feel from parents and friends that they have to do well in mathematics. This interpretation is substantiated on the negative impact of this factor for students with otherwise similar levels of socio-economic and cultural status, self-efficacy, and anxiety. There may, however, be a reversed causality in this association; somewhat weaker students may feel more pressure than their more successful peers that are otherwise similar to them on other factors.

In conclusion, our study confirmed the influences of socio-economic background on student mathematics achievement reemphasizing the need for constant and more in-depth research in this area. It seems safe to say that research on equity in education is needed also in other achievement areas. Based on comparisons with the other countries this is even more important for Slovenia due to a somewhat stronger impact of socio-economic and cultural status and a weaker mediating impact of mathematics-related self-efficacy on student achievement.

There are, of course, limitations to generalizing the results of this study. As mentioned, all data are based on students' reports. This may in-

17 Gradients in Canada and the United States are significantly smaller than in Slovenia.

fluence objectivity and comparability of data across countries as well as within. Also, it is important to be careful in assuming causality from the models. It may well be that the outcome variable - mathematics achievement - influences the levels of predictors as well. For example, evidence of high achievement naturally increases one's conviction of their capability to solve mathematics tasks. Or, parents may exert less pressure for mathematics learning when their children are high achievers. Since the model included four predictors, the observed impacts of these factors may not only or not at all be direct effects but also due to effects of possible other hidden or unmeasured variables not included in the model. In addition, we assumed only linear relationships in the model while there may be curvilinear relationships between the factors as well as with the outcome variable and additional multilevel influences. However, the findings from this model seem reasonable and informative for future methodology of the national and international educational studies as well as for educational policy.

Further studies may explore the issues addressed in this article in several directions. First, other countries may be taken into account. This could show generalizability of the present results across different cultures and educational settings. Second, the outcomes in specific mathematical sub-domains could be considered. This could show generalizability of the results across different process and content-specific achievements, which were organized in PISA in process sub-scales formulating, employing and interpreting and content sub-scales change and relationships, space and shape, quantity, and uncertainty and data. Further, with a larger number of countries or additional factors other methods could be used, like multi-level modelling to explore the proportions of between-country variance explained by the selected predictors, or structural equation modelling to explore the possible causal interrelationships of the selected factors.

References

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes.* 50, pp. 179-211.

Bandura, A. (1997) *Self-Efficacy: the Exercise of Control.* New York: Freeman.

Bandura, A. (1977) *Social Learning Theory.* New Jersey: Prentice Hall.

Bourdieu, P., and Passeron, J.C. (1990) *Reproduction in Education, Society and Culture. 2nd Ed.* London: Sage Publications.

Duru-Bellat, M. (2004) *Social Inequality at School and Educational Policies.* Paris: UNESCO International Institute for Educational Planning.

Ferla, J., Valcke, M., and Cai, Y. (2009) Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*. 19, pp. 499-505.

IEA (2014). *IDB Analyzer*. Available from: http://www.iea.nl/data.html (Accessed: 27th October 2014).

Lamb, S. and Fullarton, S. (2001) *Classroom and school factors affecting mathematics achievement: A comparative study of the US and Australia using TIMSS*. Australia: Australian Council for Educational Research. Available from: http://research.acer.edu.au/timss_monographs/10 (Accessed: 27th October 2014).

Ma, L., and Bradley, K. D. (2004) The development of mathematics achievement in secondary school individual differences and school effects. Paper presented at the Mid - Western Educational Research Association annual meeting, Columbus, OH. Available from: http://www.uky.edu/~kdbrad2/The%20Development%20of%20Math%20Achievement.pdf (Accessed: 27th October 2014).

Morony, S., Kleitman, S., Ping Lee, Y., and Stankov, L. (2012) Predicting achievement: confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research*. Available from: http://dx.doi.org/10.1016/j.ijer.2012.11.002 (Accessed: 27th October 2014).

Observatory of Economic Complexity (2014) Available from: http://atlas.media.mit.edu/profile/country/svn/ (Accessed: 27th October 2014).

OECD (forthcoming) *PISA 2012 Technical Report*. Paris: OECD Publishing.

OECD (2009) *PISA Data Analysis Manual. Spss. Second edition*. Paris: OECD Publishing.

OECD (2012) *PISA 2012 Assessment and Analytical Framework. Mathematics, reading, science, problems solving and financial literacy*. Paris: OECD Publishing.

OECD (2013a) *PISA 2012 Results: Excellence through Equity. Giving Every Student the Chance to Succeed. Volume II*. Paris: OECD.

OECD (2013b) *PISA 2012 Results: Ready to learn: Student's engagement, drive and self-beliefs. Volume III*. Paris: OECD.

OECD (2013c) *PISA 2012 Results: What students know and can do: Student performance in mathematics, reading and science. Volume I*. Paris: OECD.

National Examination Centre [Državni izpitni center] (2014) National Assessment of Knowledge. Available from: http://www.ric.si/national_assessment_of_knowledge/general_information/ (Accessed: 27th October 2014).

Pang, X., and Rogers, W.T. (2013) *Comparative Examination of the Influence of Selected Factors on Achievement in Grade 9 Academic and Applied Mathematics.* Canada: Queen's Printer for Ontario. Available from: http://www.eqao.com/Research/pdf/E/Comparative_Examination_Influence_Factors_Achievement_G9Math_en.pdf (Accessed: 27th October 2014).

Rychen, D. S., and Salganik, L. H. (eds.) (2003) *Key competencies for a successful life and well-functioning society.* Göttingen: Hogrefe and Huber Publishers.

Statistical Office of the Republic of Slovenia (2014a) SI-STAT database. Available from: http://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=0953201S&ti=&path=../Database/Dem_soc/09_izobrazevanje/07_srednjesol_izobraz/01_09532_zac_sol_leta/&lang=2. (Accessed: 27th October 2014).

Statistical Office of the Republic of Slovenia (2014b) SI-STAT database. Available from: http://pxweb.stat.si/pxweb/Dialog/varval.asp?ma=0952706S&ti=&path=../Database/Dem_soc/09_izobrazevanje/04_osnovnosol_izobraz/01_09527_zac_sol_leta/&lang=2. (Accessed: 27th October 2014).

Suan, J.S. (2014) Factors affecting underachievement in mathematics. Proceeding of the Global Summit on Education GSE 2014, 5 March 2014, Kuala Lumpur, Malaysia. Available from: http://www.kuis.edu.my/icomm/1st/e-proceedings/IC002%20FACTORS%20AFFECTING%20UNDERACHIEVEMENT%20IN%20MATHEMATICS.pdf (Accessed: 27th October 2014).

Štraus, M., and Markelj, N. (2011) Reading, mathematical and scientific literacy of students in grade 1 of upper secondary schools in Slovenia in the PISA 2009 study [Bralna, matematična in naravoslovna pismenost dijakinj in dijakov 1. letnikov srednjih šol v Sloveniji v raziskavi PISA 2009]. *Šolsko polje.* XXII (5-6), pp. 35-68. Available from: http://www.pei.si/UserFilesUpload/file/digitalna_knjiznica/SP/2011/%C5%A0olsko%20polje,%20letnik%20XXII,%20%C5%A1tevilka%205-6.pdf (Accessed: 27th October 2014).

Willms, J.D. (2003) *Ten hypotheses about socioeconomic gradients and community differences in children's developmental outcomes.* Ottawa, Ontario, Canada: Applied Research Branch of Human Resources Development Canada.

Willms, J.D. (2006) *Learning divides: ten policy questions about the performance and equity of schools and schooling systems.* Montreal: UNESCO Institute for Statistics.

Zimmerman, B.J. (2000) Self-efficacy: an essential motive to learn. *Contemporary Educational Psychology*. 25, pp. 82-91.

Zupanc, D., and Bren, M. (2010) Grade inflation in Slovenia. *Contemporary Pedagogy* [Sodobna Pedagogika]. 61 (3), pp. 208-228.

Žakelj, A., and Ivanuš Grmek, M. (2010) *Relationship between the results of national assessment of knowledge and socio-cultural background of students, school classes and homework* [Povezanost rezultatov pri nacionalnem preverjanju znanja s socialno-kulturnim okoljem učencev, poukom in domačimi nalogami]. Ljubljana: Zavod Republike Slovenije za šolstvo.

3 ABSTRACTS/POVZETKI

Abstracts/Povzetki

Darko Strajn

The PISA Syndrome: Can we Imagine Education Without Comparative Testing?

The article starts by recalling two relatively recent criticisms of PISA testing addressed to wider public: *The Guardian*, Tuesday 6 May 2014 letter addressed to PISA director Dr Schleicher and Erwin Wagenhofer's film *Alphabet*. Both of these criticisms, aimed at policy makers and even more, to the broader public, expose the dubious nature of ranking of results that inscribe PISA into the foundations of the neoliberal extension of market competition to all avenues of life. On an another level many disputes, divergent studies, books and articles predominantly in much less agitated discourse ponder the social role, impacts, advantages and shortcomings of PISA and also of other similar assessments of education, done with methods of testing, as well as rankings and benchmarking as consequences of testing. In this context Konrad Liessmann and Christian Laval criticised the "neoliberal attack on public school". Other writers expose a threat of cultural homogenization. Of course, many reflections on PISA are enunciated in the context of post-colonial studies, gender studies and other contemporary forms of critical thinking that are often associated with political anti-globalisation movements, which also include a range of alternative education practices and experiments. These criticisms cannot be easily typified, but they are mainly based on similar, albeit much more elaborated, theses as the main points of *The Guardian* letter. The deeper reasons of controversy should be seen in the paradigmatic divide, which has its roots in the gap between the continental and Anglo-American

philosophy. The outcry against PISA in *The Guardian* letter is a kind of cumulative effect of the growing bid for emancipatory education, which again strives to return to a composition of educational ideals instead of the aims comprised in more or less utilitarian and technocratic concepts of increasingly visible failure of such neoliberal projections as knowledge society, human capital, and so on. Does all this mean that such comparative testing as PISA, as its most outstanding case, becomes obsolete? In spite of all criticism, the answer should be definitely: "No!"
Key words: testing, criticism, neoliberalism, methodology, emancipation, enlightenment

Sindrom PISA: si lahko predstavljamo izobraževanje brez primerjalnega testiranja?

Članek najprej spomni na razmeroma novi kritiki testov PISA, naslovljeni na širšo javnost: pismo, objavljeno v časopisu *The Guardian*, v torek, 6. maja 2014, naslovljeno na direktorja raziskav PISA Dr. Schleicherja in film Erwina Wagenhoferja *Abeceda*. Obe kritiki, namenjeni oblikovalcem politik, in še več, širši javnosti, izpostavljata dvomljivo naravo razvrščanja rezultatov, ki vpisujeta raziskave PISA v temelje neoliberalnega posplošenja tržne konkurence v vse pore življenja. Na drugi ravni mnogo sporov, različnih študij, knjig in člankov pretežno v obliki veliko manj vznemirjenih diskurzov premišljuje o družbeni vlogi, vplivih, prednostih in pomanjkljivostih raziskav PISA in tudi o drugih podobnih evalvacijah izobraževanja z metodami testiranja, kot tudi o razvrstitvah in standardizacijah kot posledicah testiranj. V: tem kontekstu Konrad Liessman in Christian Laval kritizirata »neoliberalni napad na javno šolo«. Drugi avtorji izpostavljajo grožnjo kulturne homogenizacije. Seveda je veliko razmislekov o raziskavah PISA oblikovanih v kontekstu postkolonialnih študij, študij spola in drugih sodobnih form kritičnega mišljenja, ki se pogosto navezuje politična antiglobalizacijska gibanja, ki vsebujejo tudi vrsto alternativnih izobraževalnih praks in eksperimentov. Vseh teh kritik ni mogoče zlahka tipizirati, vendar pa v glavnem temeljijo na podobnih, četudi bolj elaboriranih, tezah kot so poglavitne točke pisma v *Guardianu*. Globlje razloge polemike je treba videti v paradigmatski delitvi, ki korenini v prepadu med kontinentalno in anglo-ameriško filozofijo. Protest proti raziskavam PISA v pismu v *Guardianu* je nekakšen kumulativni učinek vse višje stave na emancipacijsko izobraževanje, ki si spet prizadeva vrniti v nabor izobraževalnih idealov namesto ciljev izraženih v bolj ali manj utilitarističnih in tehnokratskih konceptih vse bolj vidno zgrešenih neoliberalnih projekcij družbe znanja, človeškega kapitala, itd. Ali vse to pomeni, da je takšno primerjalno testiranje kot ga izvajajo v raziskavah

PISA, odvečno in zastarelo? Kljub vsem kritikam, bi moral odgovor vsekakor biti odločni »Ne!«
Ključne besede: testiranje, kritika, neoliberalizem, metodologija, emancipacija, razsvetljenstvo

Urška Štremfel

Slovenia on its Own Way Towards Improving PISA Results

Programme for International Student Assessment (PISA) 2009 and 2012 results showed that Slovenian students performed below the Organisation for Economic Cooperation and Development (OECD) average in reading literacy. Additionally Slovenia is a European Union (EU) member state that does not successfully follow the EU benchmark, which states that by 2020, the number of low achieving students in PISA at the EU level should be less than 15%. The article discusses PISA in terms of transnational policy making and transnational problem solving. The article, using the case study of Slovenia, explains triple pressures participating countries face when performing below average in PISA comparative achievement scale (performing below international (OECD, EU) average, non-attaining of EU benchmark and common goals, non-attaining of national goals) and how these pressures are translated in the identification of policy problem at the national level and which ways participating countries have at their disposal to find the solution to the perceived policy problem. The article therefore provides policy analysis insight in the first stage of improving PISA results through the lenses of governance of problems (where PISA could be understood as international policy promotion) and policy learning theory (where PISA could be understood as instrument for lesson-drawing). In order to preserve the sovereignty of national state (Slovenia) over its educational system, the article suggests that instead of uncritical reception of international promotion of certain educational model, the more promising alternative for improving PISA results is lesson drawing. Considering lesson drawing, by providing empirical insights on the case study of Slovenia, the article shows how important it is for participating countries to have carefully defined national educational priorities and goals in order to be able to precisely define a policy problem according to its PISA results and to find a policy solution by drawing lessons from other successful participating countries.
Key words: PISA, governance of problems, lesson-drawing, low achievers, Slovenia

Slovenija na lastni poti izboljševanja rezultatov v raziskavi PISA

Objavi rezultatov raziskave Programa mednarodne primerjave dosežkov učencev (PISA) iz leta 2009 in 2012, sta pokazali, da so bili rezultati slovenskih 15-letnikov na področju bralne pismenosti primerjalno nižji kot povprečno v državah Organizacije za ekonomsko sodelovanje in razvoj (OECD) ter da Slovenija neuspešno zasleduje ciljno vrednost Evropske unije (EU), po kateri naj bi bil odstotek 15-letnikov, ki ne dosegajo temeljne ravni bralne pismenosti do leta 2020 pod 15 %. Rezultati so v slovenskem izobraževalnem prostoru sprožili vprašanja o možnih načinih izboljševanja dosežkov slovenskih učencev. Članek skozi konceptualni okvir analize politik raziskavo PISA obravnava kot obliko transnacionalnega oblikovanja politik in transnacionalnega reševanja javnopolitičnih problemov. V članku pritiske OECD, EU in nacionalnih akterjev po izboljšanju dosežkov učencev v raziskavi PISA tako osvetlimo skozi prizmo nove oblike vladavine v EU na področju izobraževalnih politik, ki raziskavo PISA razume kot primer vladavine javnopolitičnih problemov (oziroma kot primer mednarodne javnopolitične promocije) ter kot priložnost za medsebojno javnopolitično učenje med sodelujočimi državami pri izboljševanju njihovih dosežkov (oziroma kot primer učenja lekcij). Ob upoštevanju navedenih konceptov analize politik članek predstavlja uvid, kako sodelujoče države članice na podlagi podpovprečne uvrstitve v mednarodni primerjalni lestvici dosežkov PISA zaznajo javnopolitični problem (zaradi nedoseganja mednarodnega (OECD in/ali EU) povprečja, zaradi nedoseganja ciljne vrednosti EU ali zaradi nedoseganja nacionalnih ciljev na področju izobraževanja), in katere možnosti so jim na voljo pri reševanju zaznanega javnopolitičnega problema oziroma izboljšanju dosežkov učencev. V članku izpostavimo, da je za ohranjanje suverenosti nacionalne države (Slovenije) nad njenim izobraževalnim sistemom pomembno, da za izboljšanje dosežkov učencev v raziskavi PISA kritično presoja priporočila OECD kot sredstvo transnacionalne javnopolitične promocije ter premišljeno išče rešitve in dobre prakse pri drugih sodelujočih državah na podlagi učenja lekcij. Pri tem članek izpostavi pomen jasno opredeljenih nacionalnih prioritet in ciljev kot predpogoja za ohranjanje suverenosti pri iskanju lastnih rešitev zaznanega javopolitičnega problema ter identifikaciji držav pri katerih se lahko zgledujemo pri izboljševanju dosežkov naših učencev.

Ključne besede: PISA, vladavina problemov, javnopolitično učenje, nizki dosežki, Slovenija

Christine Salzer and Manfred Prenzel

Looking Back at Five Rounds of PISA: Impacts on Teaching and Learning in Germany

The German results of PISA 2012 were solid showing student performance in all domains significantly above the average of OECD countries. Nonetheless the data still point out some challenges for the next years. If PISA 2012 had been the first round of PISA, nobody in Germany would have been surprised and the overall picture would have been described as not very spectacular. However, given the history of the profound PISA-shock in 2001, the results of PISA 2012 mark a milestone of progress after twelve years of efforts to improve learning outcomes in Germany's educational context. Looking backward on PISA 2000, this paper starts with an analysis of the different aspects of the poor performance of the students in Germany at that time, including a very broad distribution and high correlation with social background and migration. The paper then discusses three major aspects of educational development: First, a thorough diagnosis of the problems in the educational system in Germany using PISA data as well as findings from other studies was important to draw adequate conclusions for measures taking into account different parts of the educational system (including e.g. pre-school or teacher training). Against this background an intense and evidence-based discourse between policy makers, researchers and the public could be started. This discourse led to a common understanding that a higher appreciation of education and educational reforms were of vital necessity. Last but not least a considerable number of nationwide, overarching programmes to improve teaching and learning with respect to educational standards was implemented. All in all the paper argues that findings from PISA have to be interpreted in the light of other types of educational research (e.g., longitudinal design, video studies). An improved public understanding of research on education helps to get acceptance for reforms. Besides political attention and engagement a strategic and systemic view is crucial to the success.
Key words: PISA, PISA-shock, Germany, improvement, education

Pogled na dosedanjih pet ciklov raziskave PISA: učinki na poučevanje in učenje v Nemčiji

Za Nemčijo so bili rezultati raziskave PISA 2012 ugodni, saj so pokazali dosežke učenk in učencev nad povprečjem držav OECD. Kljub temu pa še vedno nakazujejo nekatere izzive za naslednja leta. Če bi bil cikel PISA 2012 prvi cikel raziskave, nihče v Nemčiji ne bi bil presenečen in splošen vtis o rezultatih ne bi bil preveč spektakularen. Vendar pa glede na velik

PISA-šok v letu 2001 rezultati raziskave PISA 2012 predstavljajo prelomnico po dvanajstih letih naporov za izboljševanje dosežkov v nemškem izobraževalnem sistemu. S pogledom na raziskavo PISA 2000 v članku začnemo z analizo različnih vidikov nizkih dosežkov nemških učenk in učencev v takratni raziskavi vključujoč zelo razpršene dosežke in visoko korelacijo s socialno-ekonomskim ozadjem in priseljenskim statusom. V nadaljevanju predstavljamo tri glavne vidike razvoja izobraževanja. Poglobljena diagnoza problemov v izobraževanem sistemu v Nemčiji z uporabo podatkov raziskave PISA kot tudi drugih raziskav je bila pomembna za izpeljavo ustreznih sprememb na različnih ravneh vzgojno-izobraževalnega sistema (vključujoč na primer predšolsko vzgojo ter izobraževanje in usposabljanje učiteljev). Na teh podlagah se je lahko začela intenzivna in na podatkih temelječa razprava med oblikovalci politike, raziskovalci in splošno javnostjo. Razprava je vodila do skupnega razumevanja, da so vrednotenje znanja in prenova izobraževanja bistvenega pomena za izboljšanje stanja. Ne nazadnje pa se je začelo izvajanje vrste vsesplošnih programov za izboljšanje poučevanja in učenja na nacionalni ravni za doseganje standardov znanja. V članku zagovarjamo, da morajo biti izsledhi raziskave PISA interpretirani v luči drugih vrst edukacijskega raziskovanja (na primer, longitudinalnih raziskav, video študij). Boljše razumevanje javnosti o raziskovanju v izobraževanju pomaga pri sprejemanju nujnosti izvajanja reform. Ob pozornosti in angažiranosti politike pa sta za uspeh pomembni tudi njena strateška in sistemska naravnanost.
Ključne besede: PISA, PISA-šok, Nemčija, izboljšave, izobraževanje

Pierre Brochu

The influence of PISA on Educational Policy in Canada: Take a Deep Breath

The results from the most recent round of the Programme for International Student Assessment (PISA) revealed that while Canada remains among the top performing countries in the world, it is showing a downward trend in skills. This paper looks at how PISA results have been used since its inception in 2000 to inform education policy in a number of countries, including Canada. It summarizes the Canadian results in the global context and compares and contrasts Canada's results with those in a number of countries of interest focusing on how the initial and subsequent PISA results have been received in these countries.

In several cases, PISA was exploited to initiate new education policies, while in others it was used to justify planned or newly implemented reforms. Considering the most recent PISA results in Canada and the call

for action from several education stakeholders, this article argues that a federated country like Canada should avoid a "one-size-fits-all" approach to education reform. Furthermore, the author argues that the initial focus of PISA results on country ranking should be replaced, or at least complemented, with a look at trends over time where a country would not only judge its progress against other countries but also against itself. In addition, federal systems like Canada, where education is decentralized, offer interesting opportunities for analyzing the PISA results at a microcosmic level to study factors related to high performance not only in other countries but in other provinces, as these often share similar contexts. Among those lessons learned from PISA over the past decade, the experience in a number of countries suggests that as useful as they may be, PISA results on their own are not a sufficient basis for initiating educational reform, as the data needs to be analyzed in a context that extends beyond the assessment itself.

Key words: PISA, large-scale assessment, education policy, international comparison

Vpliv raziskave PISA na izobraževalno politiko v Kanadi: zajemite sapo

Rezultati zadnjega cikla Programa za mednarodno primerjavo dosežkov učenk in učencev (PISA) so pokazali, da Kanada sicer ostaja med državami z najvišjimi dosežki na svetu, vendar pa je trend ravni kompetenc učenk in učencev padajoč. V članku obravnavamo, kako so bili rezultati raziskave PISA uporabljeni za oblikovanje izobraževane politike v različnih državah, vključujoč Kanado, od njenega začetka v letu 2000. Rezultati za Kanado so predstavljeni v globalnem kontekstu in primerjani z rezultati v drugih relevantnih državah s fokusom na sprejemanje začetnih in nadaljnjih rezultatov raziskave v teh državah.

V več primerih so bili rezultati raziskave uporabljeni za oblikovanje novih politik, v drugih pa za utemeljevanje že načrtovanih ali na novo vpeljanih sprememb. Glede na najnovejše rezultate raziskave PISA za Kanado in zahteve po spremembah v državi s strani različnih deležnikov v izobraževanju v članku poudarjamo, da se mora zvezna država, kot je Kanada, izogniti spreminjanju izobraževanja po modelu »enako za vse«. Avtor nadalje utemeljuje, da bi morala biti začetna pozornost na razvrstitve držav nadomeščena, ali vsaj dopolnjena, s pregledom časovnih trendov, ob katerih država svojih dosežkov ne bi primerjala le z drugimi državami, ampak tudi sama s sabo. Dodatno, zvezni sistemi, kot je kanadski, kjer je izobraževanje decentralizirano, ponujajo zanimive priložnosti za analizo rezultatov PISA na mikrokozmični ravni z raziskovanjem dejavnikov,

ki se povezujejo z visokimi dosežki ne le v drugih državah, pač pa tudi v drugih provincah, saj le-te pogosto delujejo v podobnih kontekstih. Med lekcijami iz raziskave PISA v zadnjem desetletju izkušnje v drugih državah nakazujejo, da so rezultati raziskave PISA sicer uporabni, vendar sami po sebi niso zadostna osnova za odločanje in oblikovanje izobraževalnih reform, saj morajo biti analizirani v kontekstu, ki je precej širši od same raziskave.
Ključne besede: PISA, raziskave na velikih vzorcih, izobraževalna politika, mednarodne primerjave

Maria Stephens and Anindita Sen

Comparing U.S. States' Mathematics Results in PISA and Other International and National Student Assessments

In 2012, three U.S. states – Connecticut, Florida, and Massachusetts – participated in the OECD's Program for International Student Assessment as individual entities in order to obtain an international benchmark of student performance. Such subnational participation in international assessments provides value nationally by contributing to a better understanding of the variation in national statistics and, for states, by providing a sense of the global comparative health of their education systems. However, one of the challenges in using the international data is in interpreting it alongside sometimes differing data from other international and national assessment programs in which states also participate. This article thus focuses on the question: What specific factors might explain differences in the PISA 2012 mathematics results of the three U.S. participant states and their mathematics results on other recent international and national assessments? It describes the results of a comparative analysis of four possible factors: (1) differences in the overall content distribution of the items, (2) differences in relative strengths and weaknesses on content and cognitive subscales, (3) differences in sampling, and (4) differences in participating countries.
Key words: PISA, large-scale assessment, education policy, international comparison

Primerjave matematičnih dosežkov v nekaterih državah ZDA med raziskavo PISA in drugimi mednarodnimi in nacionalnimi preverjanji

Leta 2012 so tri države v ZDA – Connecticut, Florida in Massachusetts – samostojno sodelovale v OECD-jevem Programu mednarodne primerjave dosežkov učenk in učencev z namenom, da bi pridobile mednarodne

primerjave dosežkov svojih izobraževalnih sistemov. Tovrstne oblike sodelovanja enot znotraj nacionalnega izobraževalnega sistema v mednarodnih primerjavah so pomembne tudi na nacionalni ravni zaradi boljšega razumevanja različnosti v rezultatih na nacionalni ravni in, za države znotraj ZDA, ugotavljanje stanja v njihovih izobraževalnih sistemih. Vendar pa je pri uporabi podatkov raziskav med večjimi izzivi interpretacija včasih medsebojno neusklajenih rezultatov med različnimi mednarodnimi in nacionalnimi raziskavami, v katerih države sodelujejo. V članku se posvečamo vprašanju, kateri specifični faktorji lahko razložijo razlike med matematičnimi rezultati omenjenih treh držav ZDA v raziskavi PISA 2012 in drugih nedavnih mednarodnih in nacionalnih raziskavah. V članku primerjalno analiziramo štiri faktorje: (1) vsebinske razlike v razporeditvi nalog v mednarodnih preizkusih različnih raziskav, (2) razlike v relativno močnih in šibkih področjih, izkazanih na vsebinskih in procesnih podlestvicah v raziskavah, (3) razlike v metodologiji vzorčenja med raziskavami in (4) razlike v naboru in dosežkih drugih držav, ki so sodelovale v raziskavah.

Ključne besede: PISA, raziskave na velikih vzorcih, izobraževalna politika, mednarodne primerjave

Ana Kozina and Ana Mlekuž

The Predictive Power of Attribution Styles for PISA 2012 Achievement: International and National Perspective

The study explores the predictive power of attribution styles for PISA 2012 mathematics achievement from international and national perspective. For this purpose, Weiner's' attribution theory was used as a framework in explaining the differences between high and low achieving students in PISA 2012 study for both international and national data analyses. The attribution theory (Weiner, 1985, 2010) investigates the process of attributing causes for success and failure and has been widely used as a motivational framework in achievement outcomes models. In the analyses, PISA 2012 samples were used (N=309.140) in order to define the predictive value of attribution styles. In more detail, PISA 2012 measures attributional style with two question sets dealing with: (i) the measurement of attributions for failure in mathematics (constructed index FAILMAT) and (ii) the measurement of attribution for success in mathematics (a set of questions that we combined using factor analyses into two indices: internal locus of control and external locus of control (ELC)). In the analyses, we focus primarily on attribution for success in mathematics and the predictive power of newly developed indices. The national (Slovene) results are com-

pared to high and low achieving countries in European Union (Bulgaria, Romania, Estonia, Netherlands) and with international results. The results showed that attributions for success in mathematics is a significant predictor of PISA 2012 mathematics achievement in all selected countries explaining from 7 to 14 % of mathematics achievement variance (8 % in Slovenia; 11% international average). The percentages of explained variances remain high even after the inclusion of additional the index measuring attributional style in the model (FAILMAT). The students' internal locus of control significantly predicts higher mathematics achievement and external locus of control predicts lower mathematics achievement. To article ends with the implications for classroom practise being discussed.
Key words: attribution theory, locus of control, students, PISA, achievement, mathematics

Napovedna moč atribucijskih stilov za dosežke v raziskavi PISA 2012: mednarodna in nacionalna perspektiva

V prispevku na mednarodni in na nacionalni ravni ugotavljamo napovedno moč atribucijskih stilov (različnih načinov pripisovanja vzrokov uspehu in neuspehu) za dosežke iz matematike v mednarodni raziskavi PISA 2012. Teoretični okvir predstavlja v motivacijski literaturi široko sprejeta atribucijska teorija (Weiner, 1985, 2010), ki razlaga različne načine pripisovanj vzrokov uspehu in neuspehu. V analizah smo uporabili PISA 2012 podatkovne baze - mednarodno podatkovno bazo ter izbrane nacionalne podatkovne baze (N=309 140). PISA meri atribucijski stil z dvema nizoma vprašanj: (i) pripisovanje vzrokov za neuspeh (na mednarodni ravni je oblikovan v indeks FAILMAT) in (ii) pripisovanje vzrokov za uspeh. Slednji sklop vprašanj je podrobneje analiziran v prispevku. V prvem koraku smo s faktorsko analizo analizirali postavke, ki v raziskavi PISA merijo način pripisovanja vzrokov za uspeh. Na podlagi izločevalnih kriterijev smo na mednarodni ravni identificirali dva faktorja: *notranji lokus kontrole* in *zunanji lokus kontrole.* Oba indeksa sta bila kasneje uporabljena v regresijskih modelih (multipla regresija) PISA matematičnih dosežkov. Podatki Slovenije so bili primerjani z mednarodnim povprečjem ter izbranimi državami Evropske unije. Primerjalne države smo izbrali glede na njihov povprečni matematični dosežek v raziskavi PISA 2012 (dve najvišje uvrščeni državi Evropske unije: Nizozemska in Estonija ter dve najnižje uvrščeni državi Evropske unije: Romunija in Bolgarija). Rezultati kažejo, da lahko matematični dosežek v raziskavi PISA napovemo iz podatkov o načinih pripisovanj vzrokov za uspeh dijakov. Z regresijskimi modeli lahko v izbranih državah pojasnimo od 7 do 14 % variance matematičnega dosežka (8 % v Sloveniji; 11% mednarodno povprečje). Odstotki var-

iance ostanejo visoki tudi po vključitvi dodatnega indeksa atribucijskih stilov: sprejemanje odgovornosti za neuspeh pri matematiki (FAILMAT). Dijaki, ki dosegajo višje vrednosti notranjega lokusa kontrole dosegajo pomembno višje dosežke na PISA testu iz matematike. Dijaki, ki dosegajo višje vrednosti na zunanjem lokusu kontrole dosegajo pomembno nižje dosežke na PISA matematičnem testu. Na podlagi rezultatov so podane smernice za pedagoški prakso.

Ključne besede: atribucijska teorija, lokus kontrole, dijaki, PISA, dosežek, matematika

Mojca Štraus

(In)equalities in PISA 2012 Mathematics Achievement, Socio-economic Gradient and Mathematics-related Attitudes of Students in Slovenia, Canada, Germany and the United States

The study aimed at examining the roles of socio-economic background and mathematics-related attitudinal factors in explaining achievement in mathematics literacy of the PISA 2012 study for Slovenia in comparison with Germany, Canada and the United States. The data on these factors are collected through the student background questionnaires accompanying the PISA achievement tests. While (in)equalities in student achievement due to socio-economic background have long been established, it continues to remain relevant to explore to what extent motivational and attitudinal factors can mediate this influence of socio-economic and cultural status. The international context of four countries was considered. Using linear multivariate regression, the study found that while socio-economic and cultural status remains as a strong influence on achievement, students' mathematics-related self-beliefs are stronger predictors of achievement than their drive and motivation. If socio-economic and other attitudinal factors in the model are controlled, mathematics-related self-efficacy is still a strong and important predictor of mathematics achievement in all four countries. Observing students' responses to the questions about attitudes towards mathematics interesting patterns emerged between the four countries; similarities were observed between the Slovene and German students' responses as well as between the Canadian and the United States students' responses, indicating there may exist more general, for example cultural influences on these attitudes outside the educational contexts. The slopes of socio-economic gradients on mathematics achievement varied among the four countries, being relatively high in Slovenia and Germany and relatively low in Canada and the

United States. The influence of socio-economic and cultural status therefore shows the same commonalties between the four countries as the attitudinal responses. Across all four countries, the mediating impact of factors in the relationship between the socio-economic and cultural status and mathematics achievement was generally similar with exception of mathematics self-efficacy showing a somewhat different impact in Slovenia than in the other countries.
Key words: mathematics achievement, PISA, socio-economic gradient, self-efficacy

(Ne)enakosti v matematičnih dosežkih, socio-ekonomskem gradientu in stališčih do matematike v raziskavi PISA 2012 za učenke in učence v Sloveniji, Kanadi, Nemčiji in ZDA

V članku raziskujemo vloge socio-ekonomskega ozadja in stališč do matematike pri pojasnjevanju dosežkov pri matematični pismenosti v raziskavi PISA 2012 za Slovenijo v primerjavi z Nemčijo, Kanado in Združenimi državami Amerike. Ti podatki se zbirajo z vprašalniki za učenke in učence, ki spremljajo preizkuse znanja v raziskavi PISA. Medtem ko so (ne)enakosti v dosežkih učenk in učencev zaradi socio-ekonomskega ozadja že dolgo prepoznane, še vedno ostaja relevantno raziskovanje, koliko lahko motivacijski in stališčni dejavniki mediirajo ta vpliv socio-ekonomskega in kulturnega ozadja. Obravnavali smo mednarodni kontekst štirih držav. Z uporabo linearne multivariatne regresije smo ugotovili, da so poleg socio-ekonomskega in kulturnega statusa prepričanja o sebi v povezavi z matematiko močnejši napovednik dosežkov kot vztrajnost in motivacija. Ko v modelu kontroliramo socio-ekonomski in kulturni status ter druge dejavnike, so prepričanja o sebi v povezavi z matematiko še vedno močan in pomemben napovednik matematičnih dosežkov v vseh štirih državah. Analiza odgovorov učenk in učencev na vprašanja o stališčih do matematike je pokazala zanimive primerjave med štirimi državami; med seboj so si podobni odgovori učenk in učencev v Sloveniji in Nemčiji ter odgovori učenk in učencev v Kanadi in Združenih državah Amerike. To morda nakazuje obstoj splošnejših, na primer kulturnih vplivov na stališča učenk in učencev, ki lahko izvirajo izven izobraževanega konteksta. Nakloni socio-ekonomskega gradienta na matematične dosežke se med državami razlikujejo. V Sloveniji in Nemčiji je ta naklon relativno visok in v Kanadi in Združenih državah Amerike relativno nizek. Vpliv socio-ekonomskega in kulturnega statusa tako kaže podobne primerjave med štirimi državami kot odgovori učenk in učencev o stališčih do matematike. V vseh štirih državah so vplivi dejavnikov na matematične dosežke v splošnem podobni z izjemo zaznane samoučinkovitosti pri matematiki, pri kateri

v Sloveniji vpliv na dosežke nekoliko odstopa od vpliva v drugih državah.
Ključne besede: matematični dosežki, PISA, socio-ekonomski gradient, zaznana samoučinkovitost

4 REVIEW/RECENZIJA

Review/Recenzija

Slavko Gaber (ur.) (2014). *Finska v vrhu znanja 2030*. Ljubljana: CEPS.

Kadar je nekoliko obširnejši *policy paper* zelo zanimivo in kar napeto branje, ne pa dolgočasno in formalistično razpredanje ter v najhujših primerih še ideologizirano verbalno gestikuliranje, se lahko zazdi, da se »nekaj dogaja« na ravni samih paradigem, osnovnih konceptov in načinov ter shem razmišljanja, če o zadevni realnosti niti ne govorimo. Ne govorimo o besedilu, ki bi ga podpisala kaka oblastna instanca na nacionalni ali mednarodni ravni ali o dokumentu kakega strateškega foruma, ampak o prispevku finskega učiteljskega sindikata (!). CEPS (Center za študij edukacijskih strategij na Pedagoški fakulteti v Ljubljani) in SVIZ (Sindikat vzgoje, izobraževanja, znanosti in kulture Slovenije) sta zelo ažurno in požrtvovalno poskrbela za to, da je sicer drobna a zgoščena knjižica pod zgoraj navedenim naslovom in natisnjena v lični ter bralcu prijazni brošuri, dostopna zainteresiranemu slovenskemu bralstvu skupaj z lucidnim kritičnim uvodom in daljnovidno razmišljujočo spremno besedo. Posebno zanimivost besedilu knjižice daje dejstvo, da je nastalo v učiteljskih vrstah, in da ga je tudi izdal finski učiteljski sindikat, kar je že samo na sebi »inovativno«, če se izrazimo po najnovejši evropski modi.

Finska je – kot je znano – na področju izobraževanja dosegla zvezdniški status, med drugim, zahvaljujoč odličnim rezultatom v raziskavah PISA, saj je zlasti v dosedanjem 21. stoletju nenehno na ali čisto pri vrhu po izmerjenem znanju svojih učencev in dijakov. Za »češnjico na torti« pa je Finska zelo uspešna še v drugi raziskavi v okviru OECD, namreč PIAAC, ki raziskuje znanje odraslih. Ni dvoma, da je izmerjeni visoki delež znanja in sploh višina izobrazbene ravni finskega prebivalstva, integralni del tega, kar opisujemo kot

»skandinavsko« politično kulturo in kot kulturo tolerantnosti ter smisla za vzajemno sodelovanje med ljudmi. Pri vsej zavidljivi uspešnosti pa se tudi začne novi finski problem, ki so ga zaznali učitelji skupaj s svojim sindikatom in so se zato odločili, da ne bodo čakali na uradne politične instance ter so svoje analize in projekcije v prihodnost, označeno z letnico 2030, sporočili javnosti; na srečo ne samo finski ampak tudi slovenski javnosti po zaslugi že omenjenih izdajateljev knjižice in tudi po zaslugi osebnih stikov med predstavniki obeh dežel. Kot lahko razberemo iz knjižice, je problematika, ki jo je odprl finski učiteljski sindikat, precej večdimenzionalna. Visoka uspešnost finskega izobraževalnega sistema v mednarodnem prostoru je postala prvi in z notranje finske strani opazen problem. Tako uvodničar kot pisci glavnega besedila namreč ugotavljajo, da je ta uspešnost botrovala pasivnosti upravljavcev in opuščanju razmišljanja o kljub vsemu potrebnih spremembah in izboljšavah z mislijo na prihodnost izobraževanja v družbenem in ekonomskem okolju. Drugi kompleksnejši vidik, ki ga finski sindikat zaznava (ga pa spričo svoje pozicije morda ne eksplicira z vso ostrino) pa zadeva mednarodna dogajanja in predvsem pritiske, ki jih generira neoliberalna »konkurenčnost«. Sahlberg tako v svojem uvodu posebej poudarja, da poročilo vidi odvisnost prihodnosti Finske od tega, »kako dobro bo Fincem uspelo v prihodnosti zaščititi razmeroma visoko stopnjo enakosti dohodkov in pravičnost edukacije« (str. 7).

Že na prvih nekaj straneh razberemo, da so se avtorji besedila – nastalega po obširnih razpravah v letih 2012 in 2013 med članstvom – dobro zavedali, da PISA pač meri nekaj gotovo zelo pomembnih učinkov izobraževanja, a treba je razmišljati tudi o vseh drugih prispevkih in delovanjih šolskega sistema v družbi, če naj ta deluje tako, da omogoča odprte sheme družbene reprodukcije. Tu bi si dovolil pripombo, da je prav to finsko besedilo po svoji »tipologiji« vzorna reakcija na uspeh, kakor ga meri PISA. Ta raziskava je sicer na sploh v svetu deležna tudi kritičnih pripomb, katerih utemeljenost pa je treba bolj kot snovalcem in izvajalcem raziskave pripisati delovanju neoliberalne ideologije in na njej zasnovani ekonomiji. Le-to je že l. 1998 Gilles Deleuze poimenoval »ekonomija prevečne produkcije« (*surproduction*), pri čemer je v ozadju – ali nemara bolj natančno, prav na brezsramen način v ospredju, logika »brezmejne« akumulacije kapitala. Finsko besedilo tako izhodiščno opredeli »temeljne vrednote« izobraževanja in vzgoje, ki merijo tako na tradicionalne razsvetljenske vidike (veselje do učenja, pravičnost, demokracija ipd.) kot novejše vrednote (ekologija, trajnostni razvoj). V uspešni preteklosti pa se je po ugotovitvah piscev poročila zgodilo tudi nekaj poslabšanj na večini področij, od predšolske vzgoje do izobraževanja odraslih. Tudi na

Finskem je zadnja kriza udarila svoj pečat s tem, da je ogrozila stabilnost javnih financ in s tem tudi zadovoljivost financiranja izobraževanja, če o začasnem odpuščanju učiteljev in podobnih »varčevalnih« početjih niti ne govorimo.

V nadaljevanju je očitna naklonjenost finskih sindikatov javnosti izobraževanja, ki naj bi tako ostalo tudi v prihodnosti do l. 2030. Glede vloge izobraževanja v prihodnosti finski učiteljski sindikat napoveduje vrsto sprememb tako na področju gospodarstva (omejenost rasti) in na področju družbenih dejavnikov. Tako, med drugim, vidijo kot bistveno nalogo sistema edukacije, da s pomočjo »defragmentacije in integracije« deluje kot »protisila družbeni polarizaciji« (str. 27). V enem izmed poglavij tega napetega branja se ne bo odveč poučiti o spremembah, ki se v dobrem in slabem obetajo učiteljskemu poklicu. Vsekakor pa je predvidljiva potreba po večjem ugledu in samostojnosti v opravljanju poklica. Dikcija besedila v njegovih projekcijah v prihodnost je po svoje še posebej zanimiva, saj sindikat, ki ni ključni akter pri razporejanju sredstev in ni vedno povprašan glede najpomembnejših družbenih odločitev, govori o tem da »bo Finska« več vlagala v edukacijo, ki tako ne bo »strošek«, ampak investicija. Trditev v prihodnjiku je mogoče razumeti kot sugestijo in zahtevo hkrati. Besedilo se nato podrobneje ukvarja z nevarnostmi, perspektivami in možnostmi na vsakem od področij izobraževanja posebej in se pri tem izkaže s podrobnejšimi predvidevanji. To je tisti del knjižice, ki je za vse akterje v edukaciji najbolj zanimiv še posebej glede na njihovo posebno področje dela. Če bodo na Finskem drugi akterji, namreč tisti v politiki in gospodarstvu, ta predvidevanja, o katerih tu ne govorim podrobneje, vzeli dovolj resno, bo Finska nemara spet lahko vzor za ostali svet.

V teh predvidevanjih zbudi pozornost napoved ogroženosti »osnovnega izobraževanja iz umetnosti«, ki se bo po mnenju piscev umikalo iz šole, v kateri se bo število ur tega pouka zmanjšalo in se bo selilo v neformalno izobraževanje. Sindikat torej pesimistično predvideva posledice v obliki »le premožnim dostopnih« storitev; nadalje bodo učinki vidni tako v neformalnem izobraževanju kot v učiteljskem poklicu na tem področju. Prav ta vsebina besedila, ki sicer sem in tja »ne vidi« ali spregleduje omejitve neoliberalno profilirane ekonomije in ustrezajoče družbene konstrukcije, kaže na to, da se finski sindikat vendarle izrecno v polnosti zaveda, da je prav šolsko polje ključno »bojno polje« za prihodnost na sploh. Besedilo mestoma namigne na širši mednarodni kontekst, se pa ne izreče preveč na široko o tem, da je del rešitve uganke prihodnosti morda tudi v večji mednarodni sodelovalnosti. Eden od negativnih vidikov raziskav PISA, kakor se kažejo tistim, ki kritično opazujejo globalna družbena dogajanja, je namreč učinek tekmovanja med državami, s čim-

er je tudi taka »mehka« dejavnost kot je edukacija vpotegnjena v brutalna razmerja gospodarske konkurence v obliki, v kakršni jo diktira neoliberalna ideologija ter zainteresirana globalna kapitalska oligarhija. Finski sindikat se v razpravo o tem, kot rečeno, ne spušča, čeprav potrebo po njej nakaže. Pravo vprašanje o prihodnosti ne zadeva samo Finske ampak tudi vse druge dežele in misel na vzajemno pomoči in sodelovanje se ponuja kar sama.

Darko Štrajn

5 AUTHORS/AVTORJI

Authors/Avtorji

Darko Strajn, Educational Research Institute
Darko Štrajn graduated in philosophy and sociology, Faculty of Arts – University in Ljubljana. He is currently heading a programme in educational research at the Educational Research Institute. He conducts full professor lectures on film theory at the graduate School for Studies in Humanities (ISH) in Ljubljana. His research comprises of topics such as education and social change, politics, aesthetics and media.
Darko Štrajn je diplomiral iz filozofije in sociologije na Filozofski fakulteti Univerze v Ljubljani. Trenutno vodi raziskovalni program na Pedagoškem inštitutu. Kot redni profesor predava o filmski teoriji na podiplomski šoli za humanistične študije (AMEU - ISH) v Ljubljani. Njegove raziskave obsegajo teme, kot so izobraževanje in družbene spremembe, politika, estetika in mediji.

Urška Stremfel, Educational Research Institute
Urška Štremfel is a research fellow at the Educational Research Institute in Ljubljana and part-time research fellow at the Centre for Political Science Research at the Faculty of Social Sciences, University of Ljubljana. Her research interests include the European aspects of policy analysis, especially new modes of EU governance and cooperation in the field of education policy in the framework of the open method of coordination.
Urška Štremfel, doktorica politoloških znanosti, je znanstvena sodelavka na Pedagoškem inštitutu, pri svojem raziskovalnem delu pa sodeluje tudi v Centru za politološke raziskave na Fakulteti za družbene vede Univerze v Ljubljani. Njen znanstvenoraziskovalni interes predstavlja evropsko sodelovanje na

področju izobraževanja in njegov vpliv na nacionalni izobraževalni prostor. V tem okviru posebno pozornost namenja vlogi mednarodnih raziskav pri oblikovanju slovenske izobraževalne politike in izobraževalnih praks slovenskih šol ter razvoju na podatkih temelječega izobraževanja.

Christine Sälzer, Technische Universität München
Christine Sälzer is an educational researcher at the Technische Universität München (TUM), School of Education in Munich, Germany, and the Centre for International Student Assessment (ZIB). She has been the PISA National Project Manager of PISA 2012 and is currently managing PISA 2015. Her main research interests are large-scale student assessments, school absenteeism and student behaviour at school.
Christine Sälzer je raziskovalka v izobraževanju na Technische Universität München (TUM), School of Education v Nemčiji, in pri Centre for International Student Assessment (ZIB). Bila je nacionalna kooordinatorica raziskave PISA 2012 in trenutno sodeluje pri vodenju raziskave PISA 2015. Glavno področje zanimanja so raziskave dosežkov v izobraževanju na velikih vzorcih, tematsko pa odsotnost učencev od pouka in vedenjske težave v šoli.

Manfred Prenzel, Technische Universität München
Manfred Prenzel is the dean of the Technische Universität München (TUM), School of Education in Munich, Germany, and the president of the Centre for International Student Assessment (ZIB). He has been the PISA National Project Manager of PISA 2003, 2006 and 2009 and is currently managing PISA 2015. Besides his interest in Large Scale Assessments, his research deals with cognitive and motivational aspects of teaching and learning.
Manfred Prenzel je dekan na Technische Universität München (TUM), School of Education v Nemčiji in predsednik ustanove Centre for International Student Assessment (ZIB). Bil je nacionalni koordinator raziskave PISA v letih 2003, 2006 in 2009 ter trenutno sodeluje pri vodenju raziskave PISA 2015. Ob delu pri raziskavah dosežkov v izobraževanju na velikih vzorcih se ukvarja s kognitivnimi in motivacijskimi vidiki poučevanja in učenja.

Pierre Brochu, Council of Ministers of Education Canada
Pierre Brochu is Director, Learning Assessment Programs with the Council of Ministers of Education (Canada). He has been Co-National Project Manager for PISA as well as co-representative for Canada on the PISA Governing Board since 2009. He is also pursuing doctoral studies in Hu-

man Development and Applied Psychology at the Ontario Institute for Studies in Education (OISE) at the University of Toronto.
Pierre Brochu je direktor Programov raziskovanja dosežkov učenja pri Svetu ministrov za izobraževanje v Kanadi. Od leta 2009 opravlja naloge nacionalnega sokoordinatorja raziskave PISA in sopredstavnika Kanade v Mednarodnem svetu PISA. Pripravlja doktorsko disertacijo na področju razvojne in aplikativne psihologije na Univerzi v Torontu, Ontario Institute for Studies in Education (OISE).

Maria Stephens, American Institutes for Research
Maria Stephens is a senior researcher at American Institutes for Research. In this capacity, Stephens provides analytic, writing, and other technical support for projects, with two foci in recent years: on coordinating international activities in the area of assessment and indicators and on district-level education reform. Currently, Stephens manages the school review process for the Say Yes to Education effort in the Syracuse City School District and also provides support to the National Center for Education Statistics' International Activities Program. She leads activities to improve the utilization of international data at the domestic level (e.g., developing an international research database, indicator reports, and reports comparing international and domestic assessments) and assists staff in making strategic decisions regarding the future of the Program for International Assessment.
Maria Stephens je višja raziskovalka pri American Institutes for Research. V tej vlogi izvaja analitično, poročevalsko in drugi tehnično podporo projektom. V zadnjih letih dela predvsem na dveh vsebinskih področjih: na koordinaciji mednarodnih aktivnosti na področju preverjanja znanja in indikatorjev in na prenovi izobraževanja na ravni lokalne skupnosti. Trenutno vodi proces evalvacije šol za project Say Yes to Education v Syracuse City School District in izvaja podporne dejavnosti za program mednarodnih aktivnosti v National Center for Education Statistics. Vodi aktivnosti za izboljšanje uporabe mednarodnih podatkov na nacionalni ravni (npr. priprava mednarodne raziskovalne baze, poročila o indikatorjih in poročila s primerjavami rezultatov mednarodnih in nacionalnih raziskav) in sodeluje pri strateškem razvoju programa mednarodnih aktivnosti.

Anindita Sen, American Institutes for Research
Anindita Sen is a senior research analyst in the Education Department at the American Institutes for Research. Dr. Sen has worked on Annual Reports for Education data and international assessments. She has published findings from studies including the Program for International Student

Assessment (PISA), the Progress in International Reading Literacy Study (PIRLS), and the Trends in International Mathematics and Science Study (TIMSS). She has also spent many years serving as a coordinator and presenter at data training seminars and summer conferences sponsored by the National Center for Education Statistics to train advanced graduate students, university faculty, and researchers from around the country in using longitudinal and international databases in their work. Dr. Sen is a graduate of the Economics Program at the New York University, where he received both his master's degree and Ph.D.
Anindita Sen je višja analitičarka v Education Department pri American Institutes for Research. Pripravljala in analizirala je podatke za publikacije Annual Reports for Education in za mednarodne raziskave v izobraževanju. Objavila je več izsledkov iz Programa mednarodnih primerjav dosežkov učenk in učencev (PISA), Mednarodne raziskave bralne pismenosti (PIRLS) in Mednarodne raziskave trendov znanja matematike in naravoslovja (TIMSS). Vrsto let je bila koordinatorica in predavateljica na seminarjih za usposabljanje uporabnikov longitudinalnih in mednarodnih baz, ki jih je sponzoriral National Center for Education Statistics. Dr. Sen je diplomirala na Economics Program pri New York University, kjer je dosegla tudi naziva magistrica in doktorica znanosti.

Ana Kozina, Educational Research Institute
Ana Kozina is a researcher, assistant professor and a head of the Centre for evaluation studies in Educational Research Institute. Her work is in the field of developmental and educational psychology. She is focused on the developmental and time related trends of aggression and anxiety (in childhood and adolescence) their interplay and the role anxiety and aggression play on individual level, on school level and on the community level (with possible prevention and intervention designs). In the field of education she is interested in the factors related to students' achievement (school climate, social and emotional learning, motivation...). She has been involved in several national and international research and evaluation projects. Currently she is working on postdoctoral project: Development of guidelines for aggression reduction on school level based on an anxiety-aggression model and trend analyses of anxiety and aggression in Slovenia primary schools from year 2007 to year 2011. Her work is presented on national and international level (e.g. conferences, journals, monographs) on regular basis. She is a member of Editorial board: Educational research Institute Press.
Ana Kozina je diplomirana univerzitetna psihologinja, doktorica psiholoških ved in docentka za psihologijo. Zaposlena je na Pedagoškem in-

štitutu kjer je vodja Centra za evalvacijske študije. Njeno raziskovalno delo sega na področji pedagoške in razvojne psihologije. Ukvarja se z razvojem agresivnosti in anksioznosti (obdobje otroštva in mladostništva) ter njune interakcije na ravni posameznika in na ravni širšega družbenega okolja (vključno z razvojem preventivnih in intervencijskih dejavnosti). Na področju pedagoške psihologije se ukvarja s preučevanjem dejavnikov (šolska klima, socialno in čustveno učenje, motivacija ...), ki vplivajo na učne dosežke otrok in mladostnikov. Njeno raziskovalno delo vključuje vključenost v mednarodne in nacionalne raziskovalne projekte ter evalvacijske študije. Trenutno je vodja temeljnega podoktorskega raziskovalnega projekta z naslovom: *Razvoj smernic za zmanjševanje agresivnosti na ravni šol na podlagi modela povezanosti agresivnosti in anksioznosti ter analize trenda obeh pojavov v slovenskih osnovnih šolah od leta 2007 do leta 2011.* Izsledke predstavlja na nacionalni in mednarodni ravni (znanstvene konference, posveti, članki, poglavja, monografije). Je članica uredniškega odbora Založbe Pedagoškega inštituta.

Ana Mlekuž, Educational Research Institute
Ana Mlekuž holds a B. A. in political sciences, is a researcher at the Educational Research Institute in Ljubljana. She is data manager for International Civic and Citizenship Education Study (ICCS 2009) and European Survey on Language Competences (ESLC 2011) and is a co-author of several scientific articles in the field of international large scale assessments.
Ana Mlekuž, univ. dipl. pol, je zaposlena kot raziskovalka na Pedagoškem inštitutu. Je upravljavka podatkovnih baz za Mednarodno raziskavo državljanske vzgoje in izobraževanja (ICCS 2009) in Evropsko raziskavo o jezikovnih kompetencah (ESLC 2011) ter je soavtorica znanstvenih in strokovnih člankov s področja mednarodnih raziskav znanja

Mojca Straus, Educational Research Institute
Mojca Štraus is a researcher in international and national studies of different areas in education and serves as the national coordinator of the PISA Study as well as the leader of the programme of the international educational at the Educational Research Institute. In addition to research in education her research work focuses on statistical approaches to analyzing the data from international comparable studies. She leads the Educational Research Institute (Pedagoski institut) as its director.
Mojca Straus se raziskovalno ukvarja z mednarodnimi in nacionalnimi raziskavami različnih področij v izobraževanju in je nacionalna koordinatorica raziskave PISA ter vodja Programa mednarodnih raziskav v izo-

braževanju na Pedagoškem inštitutu. Njeno raziskovalno delo je poleg raziskovanja šolskega polja usmerjeno v teoretično in praktično obravnavo različnih statističnih pristopov pri analizi podatkov mednarodnih primerjalnih raziskav. Kot direktorica vodi Pedagoški inštitut.

Navodila avtorjem/-icam člankov v reviji *Šolsko polje*

Članek (praviloma v obsegu od 7000 do največ 10.000 besed) naj ima na začetku: 1) naslov ter ime in priimek avtorja/-ice, 2) povzetek v slovenskem in angleškem jeziku, do 300 do 350 besed, 3) ključne besede v slovenščini in angleščini (do 5), 4) kratko predstavitev avtorja/-ice (do 100 besed v slovenščini in angleščini), navedena naj bo tudi organizacija zaposlitve.

Prispevki naj bodo napisani v knjižni slovenščini ob upoštevanju veljavnega pravopisa, v nasprotnem primeru si uredništvo pridržuje pravico, da članka ne recenzira oziroma ga zavrne.

Če je prispevek že bil objavljen v kaki drugi reviji ali če čaka na objavo, je treba to izrecno navesti.

Prispevek naj ima dvojni medvrstični razmik, tip črk naj bo Times New Roman, velikost 12 pik (v opombah 10). Besedilo naj bo levo poravnano, strani pa zaporedno oštevilčene. Odstavki naj bodo ločeni s prazno vrstico.

Uporabiti je mogoče tri hierarhične nivoje podnaslovov, ki naj bodo oštevilčeni (uporabljajte izključno navaden slog, v prelomu bodo ravni ločene tipografsko): 1 – 1.1 – 1.1.1

Za poudarke uporabite izključno *ležeči* tisk (v primeru jezikoslovnih besedil, kjer so primeri praviloma v ležečem tisku, lahko za poudarke izjemoma uporabite polkrepki tisk). Ležeče pišite tudi besede v tujih jezikih. Raba drugih tipografskih rezov (podčrtano, velike male črke, krepko kurzivno ...) ni dovoljena. Ne uporabljajte dvojnih presledkov, prav tako ne uporabljajte preslednice za poravnavo besedila. Edina oblika odstavka, ki je dovoljena, je odstavek z levo poravnavo brez rabe tabulatorjev prve ali katerekoli druge vrstice v ostavku (ne uporabljajte sredinske, obojestranske ali desne poravnave odstavkov). Oglate oklepaje uporabljajte izključno za fonetične zapise oz. zapise izgovarjave. Tri pike so stične le, če označujejo prekinjeno bese... Pri nedokončani misli so tri pike nestične in nedeljive ... Prosimo, da izključite funkcijo deljenja besed.

Sprotne opombe naj bodo samooštevilčene (številke so levostično za besedo ali ločilom – če besedi, na katero se opomba nanaša, sledi ločilo) in uvrščene na tekočo stran besedila.

Citati v besedilu naj bodo označeni z dvojnimi, citati znotraj citatov pa z enojnimi narekovaji. Izpuste iz citatov in prilagoditve označite s tropičjem znotraj poševnic /.../. Daljše citate (več kot 5 vrstic) izločite v samostojne odstavke, ki jih od ostalega besedila ločite z izpustom vrstice in umikom v desno. Vir citata označite v okroglem oklepaju na koncu citata: (Benjamin, 1974: str. 42). Če je avtor/-ica naveden/-a v sobesedilu, priimek lahko izpustite.

V besedilu označite najprimernejša mesta za *likovno opremo* (tabele, skice, grafikone itd.) po zgledu: [Tabela 1 približno tukaj]. Posamezne enote opreme priložite vsako v posebni datoteki (v .eps, .ai, .tif ali .jpg formatu, minimalna resolucija 300 dpi). Naslov tabele je nad tabelo, naslov grafa pa pod grafom. Prostor, ki ga oprema v prispevku zasede, se šteje v obseg besedila, bodisi kot 250 besed (pol strani) ali 500 besed (cela stran).

Na vir v besedilu se sklicujte takole: (Ducrot, 1988). Stran navedka navedite za dvopičjem: (Foucault, 1991: str. 57).

Če so trije avtorji/-ice navedenega dela, navedite vse tri: Bradbury, Boyle in Morse (2002), pri večjem številu pa izpišite le prvo ime: (Taylor et al., 1978).

Dela enega avtorja/-ice, ki so izšla istega leta, med seboj ločite z dodajanjem malih črk (a, b, c itn.), stično ob letnici izida: (Bourdieu, 1996a).

Dela različnih avtorjev/-ic, ki se vsa nanašajo na isto vsebino, naštejte po abecednem redu in jih ločite s podpičjem: (Haraway, 1999; Oakley, 2005; Ramazanoglu, 2002).

Pri večkrat zaporedoma citiranih delih uporabite tole: (ibid.).

V članku uporabljena dela morajo biti po abecedi navedena na koncu, pod naslovom *Literatura*. Če so bili v prispevku uporabljeni viri, se seznam virov, pod naslovom *Viri*, uredi posebej. Če je naslovov spletnih strani več, se lahko navedejo tudi v posebnem seznamu z naslovom *Spletne strani*. Pri navedbi spletne strani se v oklepaju dopiše datum dostopa. Vsako enoto v teh seznamih zaključuje pika. Način navedbe enot je naslednji:

Knjige: Bradbury, I., Boyle, J., in Morse, A. (2002) *Scientific Principles for Physical Geographers*. Harlow: Prentice Hall.

Garber, M. (1999) *Symptoms of Culture*. Harmondsworth: Penguin.

Članki: Kerr, D. (1999b) Changing the political culture: the advisory group on education for citizenship and the teaching of democracy in schools. *Oxford Review of Education* 25 (4), str. 25–35.

Poglavja v knjigi: Walzer, M. (1992) The Civil Society Argument. V MOUFFE, Ch. (ur.). *Dimensions of Radical Democracy: Pluralism, Citizenship and Community*. London: Routledge.

Spletne strani: http://www.cahiers-pedagogiques.com/article.php3?id_article=881 (pridobljeno 5. 5. 2008).

O morebitnih drugih posebnostih se posvetujte z uredništvom.

Naslov uredništva: Šolsko polje, Mestni trg 17, 1000 Ljubljana; tel.: 01 4201 240, fax: 01 4201 266, e-pošta: info@theschoolfield.com; eva.klemencic@pei.si

Naročilo na revijo: Šolsko polje, Slovensko društvo raziskovalcev šolskega polja, Mestni trg 17, 1000 Ljubljana, e-pošta: eva.klemencic@pei.si; tel.: 01 420 12 53, fax: 01 420 12 66

Guidelines to the authors

The submission of an article to the *Solsko polje* journal should be between 7.000 to 10.000 words long. At the beginning it should include

- the author's name and address;
- a summary in both Slovene and English (from 300 to 350 words);
- 5 keywords in both Slovene and English;
- a short presentation of the author in both Slovene and English (each of up to 100 words) including his/her institutional affiliation.

The submission should be accompanied by a statement that the submission is not being considered for publication in any other journal or book collection.

The spacing of the article should be double spaced, the font Times New Roman (size 12 in the main text and size 10 in the footnotes). Paragraphs should be indicated using an empty row. There are three types of hierarchical subheadings, which should be numbered as follows:

1.

1.1

1.1.1

For emphasis, use italics only. Words in a foreign language should also be italicized. Use self-numbered footnotes.

Double quotations marks should be used for quotes in the text and single quotation marks for quotes within quotes. Longer quotations (more than 5 lines) should be extracted in separate paragraphs and separated from the rest of the text by omitting the rows and by having an indentation to the right. The source of the quotation should be in round brackets at the end of the quotation, e.g. (Benjamin, 1974, pp. 42–44).

Please mark in the text the place where a graphic product (tables, diagrams, charts, etc..) should be included, e.g. [Table 1 about here]. These products should be attached in a separate file (in 'eps,' 'ai,' 'tif' or 'jpg' format [300 dpi resolution]). The table title should be above the relevant table or the graph.

The source in the text should be referred to as follows: (Ducrot, 1988). Please quote the page for a: (Foucault, 1991, p. 57). If there are three authors, please refer as (Bradbury, Boyle and Morse, 2002) or (Taylor et al., 1978) for four or more authors.

For the works of an author that were published in the same year, distinguish between them by adding small letters (a, b, c, etc.), e.g. (Bourdieu, 1996a). Repeatedly cited works should use the following: (ibid.). Please, use the following style for each of publication:

Books:

Bradbury, I., Boyle, J., and Morse, A. (2002) *Scientific Principles for Physical Geographers*. Harlow: Prentice Hall.

Garber, M. (1999) *Symptoms of Culture*. Harmondsworth: Penguin.

Journal Articles:

Kerr, D. (1999b) Changing the political culture: the advisory group on education for citizenship and the teaching of democracy in schools. *Oxford Review of Education*. 25 (1–2), pp. 25–35.

Book chapters:

Walzer, M. (1992) The Civil Society Argument. In: Mouffe, Ch. (ed.). *Dimensions of Radical Democracy: Pluralism, Citizenship and Community*. London: Routledge.

Websites:

http://www.cahiers-pedagogiques.com/article.php3?id_article=881 (5. 5. 2008).

Šolsko polje, Mestni trg 17, 1000 Ljubljana; tel.: 01 4201 240, fax: 01 4201 266, e-pošta: info@theschoolfield.com; eva.klemencic@pei.si

Šolsko polje, Slovensko društvo raziskovalcev šolskega polja, Mestni trg 17, 1000 Ljubljana, e-pošta: eva.klemencic@pei.si; tel.: 01 420 12 53, fax: 01 420 12 66

Šolsko polje

Revija za teorijo in raziskave vzgoje in izobraževanja
Letnik XXV, številka 5–6, 2014

EDITORIAL/UVODNIK

PAPERS/RAZPRAVE

CENA: 10 EUR

ISSN 1581-6036