"Who are we writing for?": On Research Publishing in Comparative Studies based on International Large-Scale Assessments

Sverker Lindblad

Department of Education and Special Education, Faculty of Education, University of Gothenburg, Sweden
sverker.lindblad@ped.gu.se | https://orcid.org/0000-0003-3481-3686

Daniel Pettersson

Academy for Education and Economics, Department of Education, University of Gävle, Sweden
daniel.pettersson@hig.se | https://orcid.org/0000-0001-6594-6145

ABSTRACT

This study is based on an interest in interaction between science and society and how this structures science and society in tandem. In order to capture such interaction, we are analysing statements in scientific publications. The purpose of this study is to analyse relevancing in scientific publications by studying who are addressed by the research contributions and why these are considered to be relevant. Our case is the field of research labelled as International Large-Scale Assessments (ILSA), such as the OECD's Programme for International Student Assessment (PISA), created to analyse relations between educational designs and student performances.

We identified a large set of research publications by means of the search engines Web of Science and Scopus. We selected publications that were peer reviewed and based on empirical comparisons between at least two countries. A large majority were only analysing student achievement, and few were researching impacts of educational variations. Relevance statements were mostly addressing policymakers. These findings are indicating strong social structuring of much ILSA research.

KEY WORDS

international comparisons; science-society interaction; research communication; large scale assessments; educational research.

SISYPHUS

JOURNAL OF EDUCATION

VOLUME 11, ISSUE 02,

2023, PP 139-163

DOI: https://doi.org/10.25749/sis.28905

"Para Quem Estamos a Escrever?": Sobre a Publicação de Investigação em Estudos Comparativos Baseados em Avaliações Internacionais em Grande Escala

Sverker Lindblad

Department of Education and Special Education, Faculty of Education, University of Gothenburg, Suécia
sverker.lindblad@ped.gu.se | https://orcid.org/0000-0003-3481-3686

Daniel Pettersson

Academy for Education and Economics, Department of Education, University of Gävle, Suécia
daniel.pettersson@hig.se | https://orcid.org/0000-0001-6594-6145

RESUMO

Este estudo tem por base um interesse na interação entre ciência e sociedade e como isto estrutura a ciência e a sociedade em conjunto. De modo a captar tais interações, analisamos afirmações em publicações científicas. A finalidade deste estudo consiste em analisar o *relevancing* em publicações científicas através da análise dos destinatários a quem os contributos da investigação são endereçados e porque estes são considerados relevantes. O nosso caso é a área de investigação chamada ILSA – Avaliações Internacionais em Grande Escala, como o Programa Internacional de Avaliação de Alunos (PISA) da OCDE, concebido para analisar as relações entre modelos educativos e desempenhos dos estudantes.

Identificámos um grande grupo de publicações de investigações através dos motores de busca Web of Science e Scopus. Selecionámos publicações revistas por pares e baseadas em comparações empíricas entre pelo menos dois países. Uma grande maioria analisava apenas o aproveitamento dos alunos, havendo poucos a investigar os impactos de variações educativas. As afirmações de relevância eram endereçadas acima de tudo aos decisores políticos. Estes resultados indicam uma estruturação social forte de grande parte da investigação ILSA.

PALAVRAS-CHAVE

comparações internacionais; interação ciência-sociedade; comunicação da investigação; avaliações em grande escala; investigação em educação.

SISYPHUS

JOURNAL OF EDUCATION

VOLUME 11, ISSUE 02,

2023, PP 139-163

DOI: https://doi.org/10.25749/sis.28905

"¿Para Quién Escribimos?": Sobre la Publicación de Investigaciones en Estudios Comparados Basados en Evaluaciones Internacionales a Gran Escala

Sverker Lindblad

Department of Education and Special Education, Faculty of Education, University of Gothenburg, Suecia
sverker.lindblad@ped.gu.se | https://orcid.org/0000-0003-3481-3686

Daniel Pettersson

Academy for Education and Economics, Department of Education, University of Gävle, Suecia
daniel.pettersson@hig.se | https://orcid.org/0000-0001-6594-6145

RESUMEN

Este estudio se basa en un interés por la interacción entre ciencia y sociedad y cómo esto estructura la ciencia y la sociedad en conjunto. Para capturar tal interacción, analizamos declaraciones en publicaciones científicas. El propósito de este estudio es analizar *relevancing* en las publicaciones científicas estudiando a quiénes se dirigen las contribuciones de investigación y por qué se consideran relevantes. Nuestro caso es el campo de investigación denominado ILSA - Evaluaciones Internacionales a Gran Escala, como el Programa para la Evaluación Internacional de Alumnos (PISA) de la OCDE, diseñado para analizar las relaciones entre los modelos educativos y el desempeño de los estudiantes.

Identificamos un gran conjunto de publicaciones de investigación a través de los motores de búsqueda Web of Science y Scopus. Seleccionamos publicaciones revisadas por pares y basadas en comparaciones empíricas entre al menos dos países. Una gran mayoría solo analizaba el rendimiento de los estudiantes, y pocos investigaban los impactos de las variaciones educativas. Las reivindicaciones de relevancia se dirigieron sobre todo a los responsables políticos. Estos hallazgos indican una fuerte estructuración social de gran parte de la investigación de ILSA.

PALABRAS CLAVE

comparaciones internacionales; interacción ciencia-sociedad; comunicación de la investigación; evaluaciones a gran escala; investigación en educación.

SISYPHUS

JOURNAL OF EDUCATION

VOLUME 11, ISSUE 02,

2023, PP 139-163

DOI: https://doi.org/10.25749/sis.28905

“Who Are We Writing For?”: On Research Publishing in Comparative Studies Based on International Large-Scale Assessments

Sverker Lindblad[1], Daniel Pettersson

A Vignette for the Problem in Focus

> This analysis has shown that the achievement gap in reading between Indigenous and non-Indigenous students is 1.5 times larger in Australia than in New Zealand, and that this is accompanied by greater inequities in the allocation of school resources, especially shortages of teachers, in Australia. We therefore argue that education policymakers in Australia should work to ensure a more equitable allocation of school resources between Indigenous and non-Indigenous students. (Song et al., 2014, p. 195)

INTRODUCTION

The above quote presents a set of conclusions drawn from a comparative analysis of achievement gaps in Australia and New Zealand. These conclusions were addressed to policymakers telling them what to do to reduce the gaps between different categories of students. The authors emphasize and argue for the relevance of their study by showing the effects of different ways of distributing educational resources in the two countries. To our understanding such statements about relevance are indicating the intellectual organisation of a research field and its social structuration by different agents inside and outside academia (Whitley, 2000).[2] Such statements about the meaning of research contributions are here named “relevancing” which we consider as a vital ingredient in scientific conversations and in science-society interaction (Nowotny et al., 2001). Thus, analyses of relevancing practices in research publications is a way to capture such interplay and how the authors situate their contributions as being appropriate in different contexts such as research communities, policymaking, or professional decision-making (Foss-Lindblad & Lindblad, 2016). Given this, the purpose of this study is to analyse relevancing in scientific publications by studying who are addressed by the research contributions and why these contributions are considered to be relevant.

Our case is the field of research labelled as International Large-Scale Assessments (ILSA), such as the OECD's *Programme for International Student* Assessment (PISA), and the IEA research program *Trends in International Mathematics and Science Study* (TIMSS). This case is chosen since ILSA research has expanded much during the last decades and is having a huge impact on education policy discourses over the world (see for instance Grek, 2009; Lingard, 2020) and on conceptions of education and implications of education governance by the number (Lindblad et al., 2018). Analyses of ILSA relevancing are expected to inform us about how this research field is organising itself intellectually and socially. Though this research field is of a specific kind, our studies will

1 University of Gothenburg, Box 100, SE 405 30 Gothenburg, Sweden
2 This study is part of a larger project financed by the Swedish Research Council, project no. 2016-04520.

serve as a basis for discussing relations between science and society in the field of education and for knowledge organisation and academic *ethos* in the making of the complex field of educational sciences.

FRAMING THE RESEARCH PROBLEM

To start: we are interested in analysing comparative research based on international large-scale assessments to understand how this research is emerging out of the interplay between science and society. There are three different but connected aspects in framing the conceptualization of this study of ILSA research:

First, since ILSA research is highly visible in the interplay between science and society as shown in many ways [see for instance in reviews by Sellar and Lingard (2013), and Lindblad et al. (2015)], it is of special interest to analyse science-society interaction in this field of study. We turn here to Nowotny et al. (2001) who developed the concept of *agora* as a shared space for interaction between different agents in the co-production of science and society in tandem (see also Rip, 2010; Brown, 2012).[3]

Second, such tandem processes are organizing ILSA research intellectually as well as socially. Turning to Whitley (2000) we can understand educational research as a special field of the academy, like business studies and quite different from e.g., chemistry or physics. To Whitley (2000) and Engwall (1995) such a field can be labelled a fragmented adhocracy compared to the bureaucracy of physics. They argue that in physics, for instance, it is relatively easy to predict which are the most important research problems given current state of knowledge, while this is not possible in business studies (or educational sciences). In this it becomes important to acknowledge the different ways these fields are organising themselves and how these are socially structured by their fundings and different addressees. Thus, it is reasonable to expect correspondences between the social structuring and the intellectual organising of ILSA research in the making of a fragmentized adhocracy.

Third, a way to understand how ILSA research is organising itself is to analyse statements in research publications – who are the addressees, which are the research problems, and who are referred to and not. This is complementing previous studies in the history of ILSA research (Husén & Postlethwaite 1996; Lindblad et al., 2015) and the collection of different traditions in comparative education.[4] An important aspect is communication between different systems – scientific, political, or professional – where ILSA research contribution is working as a boundary object crossing system borders and by that changing context of importance for understanding their meaning (Lamont & Molnár, 2002).

MAKING EDUCATIONAL RESEARCH RELEVANT

The issues of interest to us are related to the wider problem of science-society interactions, which has been covered in many studies of higher education and societal

3 Nowotny et al. (2001) were originally interested to analyze transitions from mode 1 (science in the ebony tower) to mode 2 (science as part of democratic decisions). Instead, we are here using the concept of agora to study the social structuring of science – here ILSA research.

4 Given these characteristics of ILSA research, a Bordieuan use of the concept of field is not accurate here. A cosmopolitan market is perhaps a better metaphor.

contexts, such as Gieryn (1983), Nowotny et al. (2001) and Ziman (2002). Viewed from a perspective of global trends, issues like strategic funding, marketization and increased competition between researchers, research sectors and higher education institutions come to the fore. Discussing such issues is also vital in relation to a more general understanding of science and knowledge production in society (see for instance Furlong & Lawn, 2011). The situation in the 2010s seems to us to resemble the situation in the 1930s and 1940s and numerous debates about academic autonomy or heteronomy (that is dependency of external actors, see for instance Weber, 1946 or Gustavsson, 1971), in other words, to what extent science is governed externally or internally. However, the increased external pressures on tertiary research are, in fact, radically different. Therefore, the question about which norms educational scholars can, and should, adhere to is an important one to investigate. It has to be underlined that we are focussing on research governance, and not on research communication with a public audience (see for instance Jasanoff, 2021).

An awareness of academic heteronomy and autonomy, as well as a recognized increased external pressure on research, leads us to state that the rhetoric of relevance is important to acknowledge in debates and studies concerned with scientific norms. A well-known scientific debate in relation to this is Robert Merton's *scientific ethos* (1942/1973, 1968) and the inclusion of norms of communism, universalism, dis-interestedness, and organized scepticism (the so-called CUDOS norms), which have been interpreted as a defence of the autonomy and social value of science and its dependence on particular types of social structures. According to Merton, protecting science from external and ideological pressures by defending scientific autonomy would not be possible outside a liberal democracy – primarily consisting of a specific view of society and the collegiality of academia. Here, it is important to recognize that the defence of autonomy and the addressed incompatibility between science and politics basically concerned totalitarianism and not so much science-society or science relations in a larger perspective. However, based on these arguments, totalitarianism in, for instance, Nazism can be discussed as working in similar ways as the influence from markets and market ideologies (Kalleberg, 2007) on how a *Lingua Tertii Imperii* (Klemperer, 2006) changes the language and, as such, also changes the perception of human capacity. Merton's notions serve as a more general frame for illuminating the changing rhetoric of relevance of educational research based on Sweden as an example, as summarized in Table 1.

Table 1
Summary of periods of status, relevance, and positions of educational researchers 1900-2020

Period	Status	Type of Relevance	Researcher Position
1880-1940	Aspiring	Scientific	Postulant scientist
1940-1990	Expanding	Practical/political	State intellectual
1990-	Condensing	Contested	Academic entrepreneur

Note. Developed from Foss-Lindblad and Lindblad (2016).

In this summary, the periods are formulated as ideal types (cf. Weber, 1946, 1949) of different relevance claims in educational research over time in the case of Sweden. It is

important to note the specific case, but there might be similar changes in other national cases, especially those originally based on a disciplinary organisation of educational research (Keiner, 2019).

When educational studies were made (or re-made) into an aspiring scientific discipline, relevance was given by the fact that this discipline was identified as a scientific one. The postulant research had to adjust the demands of being scientific.

In the period following World War 2, relevance was judged by the possibility of applying research to deal with educational challenges presented by both practitioners and politicians. Educational change was here of importance to implement school reforms and to innovate educational designs and practices. Given this, important educational researchers turned into state intellectuals.[5]

At the end of the century, there were no resources for expansion, but more researchers go involved into educational issues – due to an increasing number of educated researchers and the availability of research funds. This resulted in the fact that the relevance of the results of their studies were being compared and contested. These changes were obviously accompanied by changes in the status of educational researchers who were demoted to the role of academic entrepreneurs (e.g., Elde Mølstad & Pettersson, 2019) competing for resources, which demanded a voice inside as well as outside the academy. Such changes had implications for research leading to tendencies towards decreased autonomy. This means that matters of research relevance have become problematic in current research as well as policymaking.

IDEAS ABOUT THE INTERPLAY OF RESEARCH AND SOCIETY

Statements about the relevance of science and research are frequently expressed in different research policy documents, where the authors argue about e.g., benefits of different kinds of research, or translating scientific knowledge into innovations. However, transfer from knowledge production to utilization is questioned in different discourses. We will briefly summarize some of these discourses and then turn to a specific position that we consider productive for an illumination of such issues.

The implications of governing research by various interests have been discussed for a long time, typically between two contrary positions propagating scientific autonomy or heteronomy (see e.g., Gustafsson, 1971). For instance, Weber (1946) argued for an autonomous position concerning scientific activity, where the prioritized task is to clarify relations between different facts, rather than give meaningful advice to policymakers. As such, Weber was rather sceptical about science as an activity for solving societal problems. Instead of solving such problems, he regarded that the scholars' task was recognizing phenomena, organizing them in a systematic way and trying to explain them, but not pretending to solve them. Weber explained this situation in his own words:

> Science today is a 'vocation' organized in special disciplines in the service of self-clarification and knowledge of interrelated facts. It is not the gift of grace of seers and prophets dispensing sacred values and revelations, nor does it partake of the contemplation of sages and philosophers about the meaning of the universe. (Weber, 1946, p. 152)

5 This position in Sweden was pointed out by Hunter (1994).

For Weber, the scientific vocation was completely different from that of a policymaker, which dictated a division of labour between policymakers and researchers. Such an autonomous position has been questioned on many grounds, with various alternatives presented and debated (see e.g., Nowotny et al., 2001).[6] Important in current discourses is what is labelled as useful science, for instance dealing with social challenges and effective research communication to achieve an economically responsible distribution of resources to science.[7]

Below we present a somewhat different stance – one that conceptualizes science as a communicative subsystem in society.

INTERPLAY OF SYSTEMS

We are inspired by a theoretical stance conceiving such interaction as taking place between autopoietic systems that differ in terms of modernization (Luhmann & Behnke, 1994) and globalization (Stichweh, 1996).[8] In Luhmann's theory of social systems, such differentiated functional systems have different tasks and communication systems and use different media and codes when observing their environment. For example, the system of science communicates about the production of new knowledge; a communication that is coded as a distinction between "true *versus* untrue" when referring statements to observations. In a political system the media is power, and the code is "power *versus* non-power", while in the education system media is the student and the code is "cultural competence *versus* cultural incompetence" (see e.g., Luhmann & Lenzen, 2002, or Qvortrup, 2005). Given the different media and different codes, these systems do not speak the same language. They are self-referential, but they can irritate each other and use other systems as resources and as references. A way to deal with this are to use the concept of boundary objects (Bowker & Star, 1999) and their flexibility when doing boundary crossings between systems. Research findings in ILSA research might thus be quite useful in policy discourses, for instance, but is here translated or coded in another way. Simply stated, this means that statements that are constructed in the educational research system are contested or approved by statements constructed in other systems.

Agora (see above) is the space for such interaction and contestation. But what happens there, in terms of what Nowotny et al. (2001) conceptualized as tandem processes in science and policy, or in system terms as penetrations and interpretations between autopoietic systems? Also, how do we access what is displayed in the science system to prepare interaction at the educational agora?

6 Positions of and interests in scientific heteronomy and autonomy have been debated intensively in science policy, as presented by Gustavsson (1971). This complex of problems turned into matters for sociology of knowledge (see for instance Stehr & Meja, 1984) as well as science and technology studies (STS) such as Knorr-Cetina & Mulkay (1983). For a review of different positions, see e.g., Taschwer (1996) who also deals with characteristics of STS and Luhmann's analyses of science and society. In common is e.g., their basis in a semiotic turn and emphasis on reflexivity in scientific practice, while they differ in their conceptualization of science and the Luhmann demarcation of science and non-science compared to a more fluent border in STS.

7 For instance, the formulations of EU research and innovation framework programmes are instructive in that sense. For instance, they deal with challenges, but it is necessary that the research has scientific qualities.

8 According to Luhmann (2006), a systems theory "[…] begins with a difference, the difference between system and environment […] Therefore, such theory does not begin with a unity […] Instead, it begins with a difference." (p. 38). He refers here to the thesis by De Saussure (1916) that language is the difference between different words. What the words refer to is quite another issue and is not a linguistic problem. A further theoretical stance on difference theory used by Luhmann is presented by Spencer Brown (1969) in his making a "mark of distinction" as a combination of two parts – a vertical line separating two sides and a horizontal line pointing to one of the sides and not to the other side.

THE CASE OF ILSA

In the development and organization of ILSA matters of relevance were of vital concern, for instance in the creation of the IEA - *International Association for the Evaluation of Educational Achievement* (Foss-Lindblad & Lindblad, 2016; Lawn, 2017). There, the variations in educational systems were regarded as resources – a natural laboratory in which different educational designs can be tested in practice. Lawn (2017) summarized it as follows:

> Now the world was to be conceived of as more than a meeting place: it was to be an educational laboratory where different practices in terms of school organization, curriculum content and methods of instruction were experimented with. (Lawn, 2017, p. 63)

From this point of view, the societal relevance of ILSA is of vital importance. The assessments as such were regarded as a promising way of learning about educational systems and school reforms and enabling comparisons of school performances in different educational designs and preconditions for education and schooling. These promises were materialized in an expanding number of research programmes (such as the IEA programmes named as the TIMSS - *Trends in International Mathematics and Science Study*,[9] the CIVED - *The Civic Education Study* and the ICCS - *International Civic and Citizenship Study*, [10] and the PIRLS - *Progress in International Reading Literacy Study*) [11]. Among these important research programmes, we find the OECD's PISA research programme (*Programme for International Student Assessment*).[12] These studies have expanded, especially since the 2000s (see e.g., Lindblad et al., 2015), and are often supported and organized by supranational organizations such as the OECD and the EU.

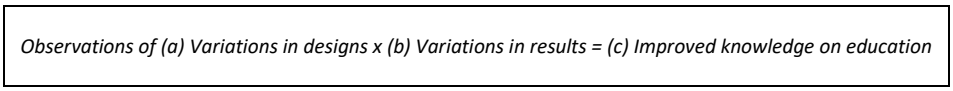
Of importance in the arguments for creating these ILSA were statements about differences in the design of education in different national contexts and that ILSA could measure the impact or efficiency of such differences in educational systems throughout the world. Thus, the IEA historically talks about its history in this way, including its nascence:

> The founders of the IEA viewed the world as a natural educational laboratory, where different school systems experiment in different ways to obtain optimal results from educating their youth. They assumed that if research could obtain evidence from across a wide range of educational systems, the variability would be sufficient to reveal important relationships that would otherwise escape detection within a single education system. They strongly rejected data-free assertions about the relative merits of various education systems and aimed to identify factors that would have meaningful and consistent influences on educational outcomes.[13]

9 https://www.iea.nl/timss
10 https://iccs.iea.nl/home.html
11 https://www.iea.nl/pirls
12 http://www.oecd.org/pisa/
13 https://www.iea.nl/brief-history-iea

In general terms, a model for such inquiries includes (a) references to differences in the ways educational systems are designed and (b) how the educational impact of such differences could be measured by international large-scale assessments to (c) improve knowledge about education. Such a model is summarized in Figure 1.

Figure 1
A general model for ILSA research and expectations of knowledge production



This model assumed that if valid observations of educational designs materialize differently in the world *and* their results are measured, then it would be possible to create new and significant knowledge about education. How such arguments are to be constructed can be observed in the research system and its internal references.

Whether this is 'true/false' is a matter of external references to the environment – in this case the education system (and its environments). Or, stated otherwise, as a matter of relating (social) facts to each other and relating ways of doing education and schooling to educational results and outcomes, both of which are conceived as social facts.[14]

As a boundary object the model for ILSA research might cross the border and move into a political system. What matters in this other system is coded as power/non-power in political terms and is relevant to successful political action – what Fuller (2016) labels as forecasting – instead of providing as valid statements as possible about a given reality that is relevant to science, for instance in relation to the model presented in figure 1.

To sum up: In this article, we inquire how this perceived 'promising way' is operated and communicated in educational research by analysing the messages on relevance stated in research publications and to whom researchers address their results and conclusions.

RESEARCH QUESTIONS

Our study of science as a communicative system is part of a larger research project funded by the Swedish Research Council and partly also based on a previous systematic review of international large-scale assessments (Lindblad et al., 2015). In these research projects we analyse the interplay between science and society at an education *agora* (cf. Nowotny et al., 2001). The agora was a square that served as a meeting-point in ancient Greece. Nowotny et al. (2001) defined it as a meeting-point where scientific, societal, and political issues could be discussed. Who meets there, and what are their agendas, positions, and resources? What do they expect from each other and how do they communicate? In the above-mentioned projects we study power and knowledge in the

14 This sentence was written in terms of a Durkheimian research programme (Durkheim, 2013); What is a social fact? In The rules of sociological method (pp. 50-59). However, we must ask how we recognize a social fact – what makes a social fact social in a social system?

field of education. The projects are preliminarily a set of studies analysing the intellectual organization and 'styles of reasoning' (cf. Hacking, 1990) of ILSA research.

In this article we only focus on research publications and on texts based on international comparative studies published in peer-reviewed journals. In this we use the term ILSA research to limit our *corpus* of articles. ILSA research should be understood as research published in peer-reviewed journals using or discussing data from international large-scale assessments like TIMSS, CIVED/ICCS and PISA for producing rhetoric on relevance concerning educational matters.

Our research questions are:

- How is the rhetoric on relevance in ILSA research formulated?
- Who do these research publications address?
- What are the implications of such rhetoric on relevance for the educational practice, policymaking, and the further production of educational knowledge?

In previous analyses of ILSA research publications (Lindblad et al., 2015) we identified different research interests and claims to answer questions expected or stated by different addressees. Referring to Weber (1946) we regard these questions being on one side related to scientific interests into what are matters of fact and how facts are related to each other, and on the other side into political interests in possibilities to achieve change and to realise visions of, for instance, good education. Given this, we ask which kinds of interests that are presented in ILSA publications and how this indicates the social and intellectual organisation of this field of study?

Following e.g., Fuller (2016) we here make a distinction between questions dealing with "how it is", for instance as how a design of an educational system actually is functioning in terms of knowledge differentiation, and "how it can be changed" as an inquiry into matters about how to reform an education system into something that does not exist (yet) but will eventually improve this system.[15] The "fact" question is based on an interest in the actual characteristics and functioning of education, while the "change" question is grounded in an interest in the possibilities to innovate or change education in certain ways.[16] To define what is possible or not is a way to achieve modal power, which is a vital aspect in policy discourses.[17]

DATA COLLECTION AND ANALYSIS

A precise set of studies based on empirical analyses in international comparisons was carried out in the following way: Using bibliometric search engines (Web of Science, Scopus, Discovery, Google Scholar) we identified many texts categorized as ILSA research

15 Fuller (2016) is here using the distinction of realist and anti-realist positions, referring to Weber and the importance to distinguish between science as a vocation and policymaking as another vocation and their respective rationalities and the importance of modal power as the capacity to decide what is and what is not possible.

16 As an example: a classical question in the making of the Swedish comprehensive school reform was research about distribution of intelligence in a cohort in relation to differentiation in educational systems. The answer from research that there was no need for an early differentiation paved the way for the undifferentiated comprehensive school reform. See e.g., Rosengren and Öhngren (1997).

17 See Fuller (2016).

from which a set of research journal articles was selected (for a detailed presentation of the search and identification of ILSA research publications, see Lindblad, Pettersson and Popkewitz, 2015). We used the Discovery search system (see Sadeh, 2013) for identifying publications and filtered the publications in various steps in order to obtain peer reviewed articles in English, presenting research based on comparisons between at least two countries. We searched separately for research based on PISA, TIMSS, PIRLS or CIVED in their variations. For a closer look, go to appendix 1.

From this *corpus* of research (for a discussion about principles for constructing a *corpus* for performing bibliometric analysis, see Mølstad et al., 2017) we selected a minor *corpus* of peer-reviewed research journal articles using or discussing data from TIMSS and/or PISA and/or CIVED/ICCS. These three domains of ILSA (TIMSS and PISA consist of several studies, and we use CIVED and ICCS as basically equivalent studies) were selected, since it is here that most of the articles (searching for the time span of 1995-2015, due to our focus of the education agora during that period) concerned with ILSA research on our selected field of interest are found (cf. Lindblad et al., 2015).

To frame our *corpus* in relation to our research interest in empirical comparative research, we limited it further to articles that were:

- presented in peer reviewed journals,
- presented results from empirical studies and
- included comparisons of two or more countries.

For example, the search for PISA publications was based on the following string *"Programme for International Student Assessment" OR "PISA" AND "education". Limited to AB Abstract and Search mode: Set for Boolean/phrase.* This search resulted in 4406 articles. Of these 2176 were excluded since they were not peer reviewed. In addition, based on abstract reading, we excluded editorials, newsletters, and articles that were not presenting research based on PISA data in two or more countries. By means of these strict inclusion/exclusion criteria we ended up with 59 articles included in our analyses. Similar search strategies were used for the two other cases with TIMSS and CIVID/ICCS.

By limiting the number of articles that explicitly compared two or more countries, the *corpus* was reduced to a level that was possible to intellectually organize and analyse. Further, after manually removing the duplicates and the misplaced articles, we ended up with a *corpus* of 135 publications (from the initial set of 8 744) that met all our criteria.

These articles were then read, mapped, and coded in terms of rhetoric on relevance, and then analysed to further our understanding of how these publications are intellectually organized and what the prevailing 'style of reasoning' is. Some of the articles use multiple ILSA for their studies and some also demonstrate multiple claims on relevance, which is important to remember when reading our conclusions.

RESULTS

IDENTIFYING STATEMENTS ON RELEVANCE AND ADDRESSEES

Messages in research publications are often attempts to communicate the relevance of a given study and its findings to other communication systems. Thus, we return to the question of which rhetoric of relevance is presented and to whom is it addressed?

Messages about relevance are often presented in the framing of a study, at the beginning of a text and/or in the final discussion about the research findings. For instance, Amanda Woods-McConney et al. (2014) presented their international study as follows:

> Given international concerns about students' pursuit (or more correctly, non-pursuit) of courses and careers in science, technology, engineering and mathematics, this study is about achieving a better understanding of factors related to high school students' engagement in science. (Woods-McConney et al., 2014, p. 1588)

The authors refer here to an internationally recognized lack of recruitment of students to scientific and technological careers, which was given as a reason for carrying out their inquiry. The data from Australia and Canada was then compared to extricate patterns of science engagement and science literacy for male and female students. The relevance of these results for science teachers is stated at the end of the article:

> (…) a major challenge for science educators is finding ways to create engaging activities within science classrooms that foster student engagement in science yet also support the development of students' scientific literacy. The evidence of this study seems to suggest that student-directed approaches in teaching and learning, and ways of broadening the view of science beyond the science classroom, beginning at home, merit further attention and research. (Woods-McConney et al., 2014, p. 1605)

The observed target of this study is therefore the education system and, more precisely, science educators.[18]

A somewhat different example is found in a publication by Pey-Yan Liou (2014). In the introduction to the text, Liou explains the relevance of the study by referring to the global economic system:

> Education not only plays an essential role in reducing people's social and economic inequality but is also the foundation of a country's economic and social development. This fact has led to the globalization of competition in almost every facet of a country's existence. Developing highly qualified human power in the fields of science, technology, and mathematics (STEM), is one the requirements to satisfy the rapid development of the global economy (…). (Liou, 2014, p. 2010)

18 The article is cited in 24 publications according to Google Scholar (2019-01-04).

In this study, Liou analysed the relations between self-concepts and science achievement based on TIMSS data and came to the following conclusion:

> The results of this study can be of valuable knowledge base for policy analysis and research regarding promoting student self-concept of learning science which is crucial to maximizing educational outcomes. (Liou, 2014, p. 2024)

The readers thus learn that the economic system is vital, and that education should produce qualified individuals for the development of the economy based on their knowledge of the sciences, mathematics, and technology. Liou also added that "(…) the results of such ILSA data are one of the most influential determinants in making educational policies in many countries" (Liou, 2014, p. 2014).[19]

These examples show two kinds of statements on relevance – one referring to the education/professional system and the other to the economic system. In the latter case, the education system is (or should be) linked to the economy. A third kind can be found in a publication by Bommi Lee (2014), who presented analyses of different designs of school tracking in relation to educational expectations in Austria (early differentiation) and Italy (late differentiation). This claim is formulated in terms of academic interests:

> Theories of social and class reproduction argue that tracking is a mechanism for the reproduction of educational and social inequality across generations. (Lee, 2014, p. 209)

Lee's conclusions read as follows:

> In this study, Italy showed significantly higher educational expectations than Austria, and this finding can be interpreted in two different ways. (…) However, careful interpretation is needed when extending the implications of the findings from this study from educational expectations to social inequalities. (Lee, 2014, p. 222)

Lee thus highlights the relevance of the findings to the scientific system, adding that the conclusions referring to the wider environment, such as the education system and society, could be problematic.[20]

Our understanding is that only the third kind of statement – by Lee (2014) corresponds to the model that was suggested when the first ILSA was developed in the 1950s (presented in Figure 1, above). This last study also included observations of the environments of the education systems in Austria and Italy, while the two previous publications only referred to very precise ILSA measurements.

We have above given examples of three kinds of rhetoric on relevance: 1) to the *educational system*, 2) to the *political system* and 3) to the *scientific system*.

Table 2 shows frequencies in making these statements in the 135 publications selected for our analysis and the lack of references to the social system. In the table we can see that some of the article's claims on relevance are double-tabulated, because they

19 This article is cited in 14 other publications, out of which six are self-citations (Google Scholar 2019-03-04).
20 The article is cited in 6 publications.

demonstrate several claims on relevance. However, this is only the case for a few of the articles; the general pattern is rather that only one claim on relevance is explicitly expressed in the articles.

Table 2

Alternative relevance claims in selected publications from the ILSA programmes PISA, TIMSS and CIVED/ICCS (1995-2015)

Relevance claims by reference to:	**Addressees:**	**Distributions of referring to programmes**			
		PISA n=59	**TIMSS** n=53	**CIVED/ICCS** n=23	**Total** n=135
Educational system	Professionals	17	12	6	35
Political system	Politicians and administrators	40	39	14	93
Scientific system	Researchers	3	2	5	10
Social system	Citizens	0	0	0	0
General information	Everybody	1	0	0	1

Note. A total of 135 articles.

As can be seen, the information presented as being of relevance to the political system predominates, followed by the educational system and its professionals. The scientific system is addressed to a little extent. To this is added that we also could identify several exceptions from this pattern. In a previous study (Lindblad et al., 2015) we analysed the dissemination of ILSA research in terms of citations within and across scientific fields. This earlier study indicated that internal communication within the scientific field to a large extent was fragmented, heterogeneous and lacked a joint fulcrum around which these kinds of studies were discussed.

The present study largely shows the same kind of pattern but enlarges our understanding of the field due to the observation that the rhetoric of relevance is only to a minor extent directed towards the scientific field, which might explain why these kinds of studies are less 'picked up' by other researchers or widely disseminated. However, as these were peer-reviewed publications they can be said to address the scientific system by default. More detailed distinctions may be needed when analysing knowledge contributions as stated in the current publications.

Of interest is the observation that none of the articles in our *corpus* address citizens. Also, it was somewhat surprising that the political system was the reference in most of the cases. These findings need to be further analysed, e.g., in comparison with other sectors in the scientific system. When reading the above table, it is also important to remember that the *corpus* only consists of articles that compare two or more nations, which may create a bias in the result and therefore warrant further investigation.

OBSERVATIONS OF ARGUMENTS PRESENTED IN RESEARCH PUBLICATIONS

The arguments put forward by ILSA research are here analysed in general terms of *explanandum* (what is explained) in relation to *explanans* (what is explaining) referring to von Wright (1971, 1983). For instance, achievement gaps (explanandum) are understood as results of differences in how schools are organized (explanans), as an example of how the general model for ILSA research as presented in figure 1 could be applied.

To illustrate this, we start with an example chosen at random. John Jerrim (2014) states that there is a growing concern "(…) that teenagers' educational and occupational plans have become detached from reality" (Jerrim, 2014, p. 197) and asks "(…) whether the educational expectations of American teenagers are indeed particularly unrealistic and poorly aligned?" (Jerrim, 2014, p. 197.) To deal with this question Jerrim turns to PISA and TIMSS data and finds that correlating student achievement test results with their educational expectations shows that low-achieving children in the US are expected to complete their college studies to a greater extent than their counterparts in other countries. Jerrim portrays this finding in diagrams and tables and discusses it in relation to social reproduction theory. He further states that expectations among American youth could be attributed to "the myth of meritocracy" (Jerrim, 2014, p. 215) and adds several alternative explanations, as summarized below:

- That the US has a large immigrant and African American population that are "(…) particularly likely to overestimate their chances of educational success" (Jerrim, 2014, p. 215).
- That an "education for all ethos" has emerged in the US where "[G]overnment policy has explicitly encouraged more and more young people to consider college" (Jerrim, 2014, p. 215).
- By tracking countries (Germany, Austria etc.) with early differentiation into vocational education tracks, and tracks towards higher education, teenagers are said to be more inclined towards vocational education.
- The rather open college education system in the US with easy entrance is said to lead to a high proportion of the entrants not graduating.

Based on these inquiries, Jerrim (2014) presents advice to policymakers in two separate directions: the first being to better prepare young people for college, and secondly to provide better help for young people to choose more appropriate educational pathways.

What have we then inferred on explanandum, explanans and claims on relevance based on our observations of this publication? Firstly, that explanandum is more a question of the unrealistic expectations of higher education amongst US youth. A first ambition is to test whether this is a 'fact' and then, according to a set of descriptive analyses, conclude that this can either be regarded as valid statement, or deviates somewhat from the normal compared to other countries.[21] Given this, statements about the *unrealistic expectations of higher education among US students* are then regarded as the explanandum. But what is the explanans in this example? Although it is difficult to

21 Jerrim (2014) is rather cautious here, but we think this can be regarded as a valid description according to the analyses carried out.

observe what the explanation(s) are based on the statements in the current article, we can identify a predominantly statistical style of reasoning (cf. Hacking, 1990), correlations are not explanations in the strict sense of the term, and neither variation over countries nor population taxonomies in terms of unrealistic expectations work as explanations. According to our understanding, an educated guess about the impact of late differentiation on education expectations could be explored and tested in accordance with the ILSA research model. To us the argument presented is based on a rather weak connection between explanandum and explanans that avoids the complexity of the problem and has littler of explanatory power.[22]

But how about explanandum and explanans in the analysed ILSA studies in more general terms? We attempt to answer this question below.

Explanandum - what is to be explained in the articles in our corpus is predominantly perceived as achievement gaps measured by performance, in e.g., science, mathematics and reading, which are then correlated by pointing to patterns of social and psychological factors, such as student careers, family status, college attendance or dropouts. We identified three dominants, sometimes interrelated, research problems:

- *Equity Problems*: Are there biases in education for e.g., gender, social class, or ethnicity and, if so, why do these occur? This problem is often connected with the search for imperatives to increase equity in education.
- *Efficiency problems*: Are there differences in performances between education systems or schools? Here we find e.g., differences in performance between countries or analyses of measures that are assumed to improve efficiency, such as tracking, or teacher performance pay.
- *Direction problems*: How should student expectations or engagement in different subjects and careers be developed, for instance in relation to science studies and occupations?

Considering *explanans*, we identified the following kinds of analyses in relation to the data used- in the research publications:

- *Internal explanans*, where different variables in the same dataset (that is for a research programme) are related to each other, e.g., student homework and reading performance.
- *External explanans*, where different data sets are combined, e.g., to capture language comprehension over time. The point is that different measurements are related to each other to develop patterns or regularities. Another external explanans in our corpus is based on the identification of population groups, e.g., income distribution, indicators of gender equity or kinds of migration that can be related to different achievement gaps.
- The logic or principles that relate internal and external explanans are specified, tacitly or explicitly through an abstract model of educational design by thinking in terms of a 'system'. The educational 'system', as a model for analysis, is used to capture student performances or achievement gaps. Thinking in terms of the design of the 'system' is about administrative characteristics of education, e.g., in terms of elements that

22 This study is cited 14 times (2019-02-12), which is in the middle in terms of this measure.

can be managed through organizational theories about, for example, tracking and differentiation.

- The last kind of explanans that we found in our corpus refers to improving school results through interventions in the organization and management of schools with a focus on, e.g., resource distribution or teaching.

These different kinds of explanans are normally combined in different ways to improve the explanandum, e.g., by discussing taxonomic groups, system organization and so forth, which means that we must describe and analyse such combinations.

In sum, ILSA research is here presented as a specific kind of collection, analysis, interpretation, and presentation of data. The collected data consists of information about individuals in different contexts, statements about their individual characteristics and performances in different kinds of tests, combined with information about their education and schooling. The data is analysed by means of different operations and the patterns achieved are interpreted in educational terms, as shown below. This is a specific kind of production of statements and of valuating objectivity within the framework of this style of reasoning. Here it is important to acknowledge that in this research there are limits that are interpreted as going from correlations to causalities related to differences in indicators and what is indicated. Ways of dealing with these limits are important when assessing the quality of ILSA research.

Arguments in terms of explanandum and explanans in the analysed ILSA publications have previously been used and analysed by Lindblad et al. (2015). Based on this review, a conclusion was that:

> A broad result concerns what to be explained – i.e., the explananda. We noted in the reviewed articles a very large share of identifications of achievement gaps over population taxonomies – e.g., classifications in terms of social class or gender – pointing to inequities and how these inequities were associated with different kinds of education measures or to contextual variations – e.g., gender inequity coefficients in different countries. To a much lesser extent differences in efficiency were analysed – e.g., school performances over educational measures or school systems. (Lindblad et al., 2015, p. 147)

Frequently, explanandum and explanans were based on observations of the same – or similar – data from the same research programme. To some extent they were derived from different sources (for an example of this infrequent operation, see Lee, 2014). Based on information derived from our coding of the 135 articles in our corpus, the following observations can be made: Firstly, only a small proportion refer to something outside the ILSA programmes (such as references to education systems). Secondly, the analyses are primarily based on correlations between different categories and/or variables within the same set of observations. And thirdly, the results are descriptive or based on explorations – and here we observe a lower degree of hypothesis-testing than expected.

Table 3
Observations in ILSA publications

Aspects	Alternatives	PISA (n:59)	TIMSS (n:53)	CIVED/ICCS (n:23)	Total 135
What to be inquired?	Student achievement	49	46	16	111
	School characters	2	2	6	10
	Education system characters	8	5	1	14
How to be inquired?	Formal (only ILSA)	46	43	20	109
	Formal plus formal (ILSA in combination with other LSA)	5	5	0	10
	Inside – outside (using ILSA for inquiring other social phenomena) and substantial facts	8	5	3	16
Type of account?	Descriptive	7	5	0	12
	Explorative – hypothesis-generating	44	23	15	82
	Hypothesis-testing	2	25	8	41

Note. Numbers over publications. Some publications are double tabulated in the table when several alternatives are evident in the same publication.

The Jerrim (2014) study presented above fits rather well with the dominant outcomes in Table 3. What is to be inquired is based on student achievements, while how to inquire means analysing patterns based on formal analyses (ILSA plus other large-scale assessments) in a hypothesis-generating account. To this is added our frequent observations of weak connections relating explanandum and explanans to each other in several empirical comparative papers in terms of measurements that lack explanation, e.g., in terms of casual mechanisms or well-grounded theoretical statements. In part, this is a consequence of the constraints in making valid statements according to statistical thought styles (for a principal elaboration, see Hacking, 1990). These observations of arguments often occur in combination with observations of what we consider as abstract, or not very distinct, statements of relevance when addressing policymakers, for instance that they should consider the results presented. Our more extensive analyses (e.g., in Lindblad et al., 2015) suggest that such weak connections and abstract reasoning frequently occur in the ILSA international comparative research publications analysed here.

CONCLUSIONS

In this study we have identified and analysed a set of ILSA research publications and their rhetoric on relevance in terms of which system they are referring to and who they are

addressing. Mostly, the authors claim that their research is of relevance for the political system – in terms of insights into education matters, but also as directives about what to do with educational challenges, such as diminishing gaps between different taxonomic categories in the population to be educated or increasing the efficiency of education. It can therefore be suggested that research relevance is mostly presented in instrumental or technological ways – as producing contributions of importance for political or professional insights and actions. In a vocabulary suggested by Bunge (1966), who argues that it is vital to differentiate between the search for success and the search for truth, the latter ambition is less frequent. A couple of more specific comments are also possible in this context.

Firstly, the information in the analysed *corpus* of articles often displays rather weak connections between explanandum and explanans, be it patterns of correlations or difference in means between different taxonomic groups. One reason for this is the design of ILSA based on one-shot observations (e.g., no longitudinal data on performances or no controlled variations in terms of experiments etc.). Another reason is the rather infrequent testing of hypotheses (e.g., on variations of educational design and achievement gaps over student SES).[23] A consequence of this that needs to be further tested is whether these knowledge contributions are considered significant inside the specific research system or not. This suggestion can be checked by analysing how this kind of research is received in the education research system, and specifically in the ILSA research sub-system. Which knowledge contributions are received, which arguments are formulated, and what are their significance? Another implication might be that such in-distinct information does not function as an 'irritating function' for the systems addressed, so that a recommendation based on a weak correlation would most likely not lead to policy 'hiccups' or professional reorientations in action.

A second preliminary conclusion, based on the first, is that the original idea of ILSA as 'a world-laboratory' does not seem to be working in the sphere of published research publications at an elevated level. ILSA publications are mostly statements about measurements of indicators and patterns of indicators, with little or no statements about variations in educational designs. To our understanding, this is a matter of incomplete argumentation with only one kind of observation of the system of education – referring to ILSA as such (it should be added that some examples of ILSA research do put forward strict and theoretically interesting conclusions and relate different facts to each other in arguments, although this is not the observed rule). However, from these conclusions it does not follow that the ILSA idea of knowledge production is falsified. We have identified several high-quality research publications based on interesting and valid results and acknowledge the potential of strict and productive ILSA research; an issue that needs to be further elaborated on.

Finally, in this article we have focused on how ILSA in research papers makes itself relevant at the education agora where different communicative systems meet. In terms of relevance, we note that these articles are mostly directed towards political actors and educational systems. When analysing arguments and possible directives based on these observations, we conclude that they are not very powerful as such. However, based on our own and other studies (see for instance Carvalho, 2012; Martens & Niemann, 2013; Sellar & Lingard, 2013), we can also conclude that this kind of research is highly visible in policy discourses. A hypothesis is thus that ILSA research works as a flexible boundary object at the education agora and is used for cherry-picking as well as the omission of convenient ILSA research results. Given this, it is reasonable to state that ILSA research often has an indistinct, yet very usable, position at the educational agora, compared to other agents' stronger references to significant (social) facts and other interests in education.

23 For an exception, see the Lee (2014) study presented above.

AUTHORS CONTRIBUTION

Both authors have contributed equally in all processes of writing the article.

REFERENCES

Bowker, G., & Star, S. L., (1999). *Sorting things out. Classification an its consequences*. The MIT Press.

Brown, C. (2012). The policy agora: how the epistemological and ideological preferences of policy-makers affect the development of government policy. *Human Welfare*, *1*(1), 57-70.

Bunge, M. (1966). Technology as applied science. In *Contributions to a Philosophy of Technology* (pp. 19-39). Springer.

Carvalho, L. M. (2012). The fabrications and travels of a knowledge-policy instrument. *European Educational Research Journal*, *11*(2), 172-188.

De Saussure, F. (1916/1959). *Nature of the linguistic sign. Course in general linguistics*. Philosophical Library.

Durkheim, E. (2013). *The evolution of educational thought: Lectures on the formation and development of secondary education in France*. Routledge.

Elde Mølstad, C., & Pettersson, D. (2019). *New Practices of Comparison, Quantification and Expertise in Education: Conducting Empirically Based Research*. Routledge.

Engwall, Lars (1995). Management research: A fragmented adhocracy? *Scandinavian Journal of Management, 11(3*), 225-235.

Foss-Lindblad, R. F., & Lindblad, S., (2016). On relevance and norms of science in times of restructuring: Educational research in Sweden. In *Political pressures on educational and social research* (pp. 66-76). Routledge.

Fuller, S. (2016). *Post Truth: Knowledge as a Power Game*. Anthem Press.

Furlong, J., & Lawn, M. (2011). *Disciplines of Education: Their Role in the Future of Education Research*. Routledge.

Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American sociological review*, *48*(6), 781-795.

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

Gustavsson, S. (1971). *Debatten om forskningen och samhället: en studie i några teoretiska inlägg under 1900-talet* (Doctoral dissertation). Acta Universitatis Upsaliensis.

Hacking, I. (1990). *The Taming of Chance.* Cambridge University Press.

Hunter, I. (1994). *Rethinking the School: Subjectivity, Bureaucracy, Criticism*. Routledge.

Husén, T., & Postlethwaite, T. N. (1996). A brief history of the international association for the evaluation of educational achievement (TEA). *Assessment in Education: principles, policy & practice*, *3*(2), 129-141.

Jasanoff, S. (2021). Knowledge for a just climate. *Climatic Change, 169*(36). https://doi.org/10.1007/s10584-021-03275-x

Jerrim, J. (2014). The Unrealistic Educational Expectations of High School Pupils: Is America Exceptional?. *The Sociological Quarterly*, *55*(1), 196-231.

Kalleberg, R. (2007). A reconstruction of the ethos of science. *Journal of Classical Sociology*, *7*(2), 137-160.

Keiner, E. (2019). 'Rigour','discipline'and the 'systematic': The cultural construction of educational research identities?. *European Educational Research Journal*, *18*(5), 527-545.

Klemperer, V. (2006). *The Language of the Third Reich*. Continuum.

Knorr-Cetina, K. D., & Mulkay, M. (1983). *Science observed: Perspectives on the social study of science*. Sage.

Lamont, M., & Molnár, V. (2002). The study of boundaries in the social sciences. *Annual review of sociology*, *28*(1), 167-195.

Lawn, M. (2017). Europeanizing through expertise: From scientific laboratory to the governing of education. In *World Yearbook of Education 2018* (pp. 53-69). Routledge.

Lee, B. (2014). The influence of school tracking systems on educational expectations: a comparative study of Austria and Italy. *Comparative education*, *50*(2), 206-228.

Lindblad, S., Pettersson, D., & Popkewitz, T. S. (2015). *International comparisons of school results: A systematic review of research on large-scale assessments in education*. The Swedish Research Council.

Lindblad, S., Pettersson, D., & Popkewitz, T. S. (2018). *Education by the Numbers and the Making of Society: The Expertise of International Assessments*. Routledge.

Lingard, B. (2020). Enactments and resistances to globalizing testing regimes and performance-based accountability in the USA. In *World Yearbook of Education 2021* (pp. 279-293). Routledge.

Liou, P. Y. (2014). Examining the Big-Fish-Little-Pond Effect on students' self-concept of learning science in Taiwan based on the TIMSS databases. *International Journal of Science Education*, *36*(12), 2009-2028.

Luhmann, N. (2006). System as difference. *Organization*, *13*(1), 37-57.

Luhmann, N., & Behnke, K. (1994). The modernity of science. *New German Critique*, (61), 9-23.

Luhmann, N., & Lenzen, D. (2002). *Das Erziehungssystem der Gesellschaft*. (1. Aufl.). Suhrkamp.

Martens, K., & Niemann, D. (2013). When do numbers count? The differential impact of the PISA rating and ranking on education policy in Germany and the US. *German Politics*, *22*(3), 314-332.

Merton, R. K. (1942/1973). *The sociology of science: Theoretical and empirical investigations*. The University of Chicago Press.

Merton, R. K. (1968). *Social theory and social structure*. Simon and Schuster.

Mølstad, C., Pettersson, D., & Forsberg, E. (2017). A Game of Thrones: Organising and Legitimasing Knowledge through PISA-research. *European Educational Research Journal, 16*(6), 869-884.

Nowotny, H., Scott, P., & Gibbons, M. (2001). *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*. Polity Press.

Qvortrup, L. (2005, December). Society's Educational System. In *Seminar. Net*, 1(1).s

Rip, A. (2010). Social robustness and the mode 2 diagnosis. *Science, Technology & Innovation Studies*, 6(1) 71-74.

Rosengren, K. E., & Öhngren, B. (1997). *An Evaluation of Swedish Research in education*. Vetenskapsrådet.

Sadeh, T. (2013). *From Search to Discovery*. IFLA WLIC.

Sellar, S., & Lingard, B. (2013). *PISA and the expanding role of the OECD in global educational governance*. Symposium Books.

Song, S., Perry, L. B., & McConney, A. (2014). Explaining the achievement gap between Indigenous and non-Indigenous students: An analysis of PISA 2009 results for Australia and New Zealand. *Educational Research and Evaluation*, *20*(3), 178-198.

Spencer Brown, G. (1969). *Laws of for*. Allen and Unvvin.

Stehr, N., & Meja, V. (Eds.). (1984). *Society and Knowledge: contemporary perspectives in the Sociology of Knowledge*. Transaction Books.

Stichweh, R. (1996). Science in the system of world society. *Social Science Information*, *35*(2), 327-340.

Taschwer, K. (1996). Science as system vs. science as practice: Luhmann's sociology of science and recent approaches in science and technology studies (STS)—a fragmentary confrontation. *Social Science Information*, *35*(2), 215-232.

von Wright, G. H. (1971). *Explanation and understanding*. Cornell University Press.

von Wright, G. H. (1983). *Practical Reason.* Basil Blackwell.

Weber, M. (1949). *Max Weber on the methodology of the social sciences*. Free Press.

Weber, M. (1958). Science as a Vocation. Daedalus, 87(1), 111-134.

Whitley, R. (2000). *The Intellectual and Social Organization of the Sciences*. Oxford University Press.

Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2014). Science engagement and literacy: A retrospective analysis for students in Canada and Australia. *International Journal of Science Education*, *36*(10), 1588-1608.

Ziman, J. (2002). *Real science, what it is and what it means.* Cambridge University Press.

*

Received: December 1, 2022
Revisions Required: May 8, 2023
Accepted: June 12, 2023
Published online: June 30, 2023

APPENDIX 1

IDENTIFICATION AND FILTRATION OF ILSA RESEARCH

Steps in filtering	CIVED and ICCS	PISA	TIMSS	Total
Hits in Discovery	947	4406	3391	8744
In peer-reviewed journals only	658	2230	1879	4786
Publication date (differs for the separate ILSA)	653	2225	1877	
Limit to Academic journals	647	2179	1829	
Limit by subject	406	1811	1553	
Limit to English	384	1662	1507	
Limit by publication	215	849	766	
Removal of duplicates by Discovery *	82 (+16)	357	317	
Removal of duplicates based on abstract	93	333	299	
Manual removal of Newsletters	93	258	265	
Manual removals of editorial articles	92	248	259	
Manual removal of misplaced articles based on abstracts	54	218	251	
Manual classification of the abstracts into A, B and C (see table 6 below for the result)	43	138	182	
Manual re-classification after reading available articles	40	131	182	353
Available articles	39	106	140	
Manual classification only in two or more countries	23	59	53	135

* Note that in the case of CIVED and ICCS a complementary Discovery search has been carried out finding 16 more articles included in the study.