The Translation and Circulation of "Evidence": Obviousness, Demonstration, and Effect in Educational Research in Germany and England

Susann Hofbauer

Peter Kelly

Anna Beck

Abstract

In current educational research communications, especially in English and increasingly also in German publications, the term "evidence" refers to international homogenizing gold standards and is often linked to European evidence policies, large scale assessments and justified and proven knowledge. Against the background of international communication and the related circulation of terms and concepts, this paper analyses German and English OECD publications that recommend the development of educational research in 1970s and 90s. In England, the OECD and thus the external perception of their own education system hardly played a significant role due to years of awareness of the need for reform. Rather, Hargreaves's lecture (1996) about the disappointing effects of educational research when compared with the achievements of evidence-based medicine was decisive for the evidence movement. In addition, non-university institutions have gained legitimacy through acting as "evidence" providers for the school system. Based on analyses of German educational research literature, we show that "evidence" appeared in neurological, medical, technological, and economic texts in the 90s and early 2000s. Usage of "evidence" increased after the PISA shock in 2000 and is now linked to the expression of disciplinary development into empirical educational research. However, based on relatively stable patterns of communication and interpretation in two academic cultures, it is shown that the connectivity of an epistemological term like "evidence" does not necessarily reduce misunderstanding.

1. Introduction

Through its inflationary use in current national as well international agendas, political and scientific communications within the testing and standardization practices of the Global Educational Reform Movement in English-speaking countries (Sahlberg in Fuller & Stevenson, 2019), "evidence" can be seen as an empty, floating signifier (Krejsler, 2017). By mobilizing knowledge in the other social fields, it joins concepts, constructions, and efforts to address the research-practice gap (Cooper, Levin, & Campbell, 2009). Simultaneously it can be discussed within the research community regarding the evaluation of research criteria, which allows research to claim to be "useful" or "evident" in the first place (Gough, 2021; Tseng & Nutley, 2014, critics see Biesta, 2010). Currently, the relevance-influencing rhetoric around evidence-baseness is so dominant in policy making on teacher education that Helgetun and Menter (2022) paraphrase this as a *central rationalized myth*. In Germany, too, the use of the term "*Evidenz*" is becoming more frequent. Over the past two decades, and

especially after the PISA shock, science policy supported the development of empirical educational research[1] and evidence-production more intensively (Zapp & Powell 2016). The establishment of empirical educational research and large-scale assessments also led to a change towards a more international academic communication. Due to the attention-drawing effect of the 2000 PISA shock, however, it is often overlooked that before and during PISA there were already recommendations by the *Organisation for Economic Co-operation and Development* (OECD) for Germany as well as for other nations when it comes to educational research. However, PISA and the OECD have had less impact in England because it has an established tradition of interconnecting knowledge production, educational politics, and the needs of the profession. Before 2009, there was relatively little reporting or public interest (Martens & Niemann, 2013) in this area, and it was only after 2010 that policy was justified, sometimes retrospectively, as a response to PISA results (Pollitt & Bouckaert, 2011). Grey (2020) identifies how the OECD has been largely excluded from making recommendations or offering policy advice in England.

Currently, "evidence" is linked to scientific gold standards that divide knowledge claims qua method, rationale, and reliable findings into evident and non-evident knowledge. It serves as orientation, demarcation but also as a communicative, international reference within educational research. Epistemological concepts and terms seem to be

globally subject to all scientific disciplines through the claim of "normal science" to attain scientific truth (Kuhn, 1962). Thus, terms that circulate internationally as scientific standards like "rigor" (*Strenge, Präzision, Stringenz, Härte, Unerbitterlichkeit*), "discipline" (*Disziplin, Fach- und Lehrrichtung, Benehmen, Strafe*) or "systematic" (*systematisch, methodisch, gründlich, planmäßig, zielbewusst*) are often translated directly. They construct educational research identities but still they do not have the same meaning, depending on the academic, historical, and social context (Keiner 2019). The concept of "evidence," which is regularly translated as "*Evidenz*" in German, suggest a certain clarity and common sense in what is being talked about.

In showing the facets of the concept of evidence, its emergence, its (inter)connotations, and its use, we will proceed in an exploratory manner. In doing so, we compare English and German reports of the OECD regarding their references and translations, describing the "evidence" discussions in England and analyzing educational research literature using "*Evidenz*" as a term. In doing so we ask if "evidence" and "*Evidenz*" have an identical meaning in different educational research cultures and contexts? How did the terms "evidence" and "*Evidenz*" circulate in the OECD recommendations? How did the term emerge in German educational research?

In a *first step*, we discuss the theoretical frame which refers to the history of knowledge as well as to Bourdieu's conditions of the

---

[1] In the German-speaking area, however, educational research tends to be a singular, uniform circumscription with its own denominations, publication formats, and institutional anchoring in the German Educational Research Association (GERA), but with strong internal differentiation. In English, it tends to be the plural form, i.e., the "foundation disciplines" (sociology, psychology, philosophy and history of education). When we speak of a more general understanding of the sciences of education and *Bildung*, we use the term

educational research in the sense of international usage. Precisely because of the interdisciplinary character in methods, thematic specialisations, and theories, we differentiate between *empirical educational research* that prefers empirical, rather quantitative methods with focus on educational institutions, especially schools (Lawn & Furlong, 2007). This research direction distinguishes itself from non-methodological science. By refereeing to a more humanistic approach, we use *science of education*.

circulation of ideas (2). In a *second step*, we reconstruct the movement of "evidence" on the level of European educational policy as well as the German and English educational research level. The longer academic "evidence" movement in the United Kingdom and particularly England can be seen in the evidence-based or evidence-informed policy and practice reform agenda since the 1980s and the creation of knowledge transfer institutions.[2] In Germany, the concept of evidence-based research has increasingly emerged right after the period referencing neurological knowledge in the 90s and attention to the *Programme for International Student Assessment* (PISA) after 2000 (3). In the last step we suggest "evidence" has a methodological, disciplinary-historical as well as an inner communication side and a side directed towards external representation, persuasion, and effect (4).

2.   The Translation and Circulation of the Obvious

Theoretically, two approaches are meaningful for our work, one that considers the circulation of knowledge historically and one that understands it in a field-specific way. Both intersect in the study of social, in this case scientific and political production and circulation of knowledge. The former considers the circulation of knowledge without hierarchies, but between different social spheres and institutions considered as fields of knowledge, which sets, forms and take up knowledge impulses. In the history of knowledge, the differentiation of scientific disciplines, scientists, their texts, concepts, belief systems or styles of thinking (Fleck, 1979), methodological and

theoretical approaches are of particular importance (Sarasin, 2011). The history of the concept of evidence can be dated back to the 17th and 18th centuries. The facets of seeing and proving that are expressed, reach back to the *epoch of evidence* that Campe (2015) situates in the period between Descartes and Kant. In this time, questions were raised about the conditions and rules under which cognition and scientific knowledge were possible. In addition to formal language, diagrammatic and pictorial symbols gained importance as methodological evidence, which clarify convictions through evidence, where images convey the "obvious" – something-before-your-eyes (*Ersichtlichkeit - Visibility, Offensichtlichkeit - Obviousness, Augenscheinlichkeit - Appearance*). "In this way *Darstellung* receives its modern primacy in knowledge: only what is capable of being represented counts as a possible object of knowledge" (Campe, 2010, p. 287). He therefore names this time as the epoch of *representation*, too. Moreover, the term also circulated between disciplines and was an expression of scientific community efforts to improve disciplinary knowledge production, mechanisms of proof, and knowledge persuasion (see Chandler, Davidson, & Harootunian 1994).

Translations and *circulation* are essential components of scientific communication and therefore scientific knowledge. Only through publication and the accompanying circulation - i.e., the persuasion of knowledge claims, their reception and dissemination – can knowledge appear as knowledge at all. However, scientific texts and ideas can circulate without context, because the

---

[2] It also features hidden lines of discourse that pertain to the *teacher-as-researcher* approach, meaning individual-based, school-based case studies rather than large-scale comparative studies. We do not go into this in detail, but it can be discerned in the figures of reasoning why research is a necessary part of successful practice and decision making

and why teachers are stakeholder of reform and therefore should engage in research. However, in this cycle it also addresses justifications of the "nature" of change and research (Hammersley 2007; Colucci-Gray, Das, Gray, Robson, & Spratt, 2013).

socio-cultural and historical conditionality of knowledge production is usually not discussed in the publication. The second approach by Pierre Bourdieu views the notion of a context-free circulation of scientific ideas critically. Recipients interpret, recontextualize and use the ideas and knowledge offered in connection with their socio-culturally embedded field of expertise. Bourdieu sought to examine the "space of theory" (*Theorieraum*) – an awareness of analytically controlling one's own scientific work (Jurt, 2014, p. 159). To avoid misunderstanding and misconceptions in the reception of ideas, knowledge and concepts, Bourdieu points out the necessary national or historical conditionality, the locally given problematics and categories of perception and evaluation (Bourdieu, 2002a). He thus opposed the "denial of an external reference" and cultural and scientific goods as "self-sufficient realities" (Bourdieu, 2002b, p. 187). Furthermore, Bourdieu views *translations* less in a linguistic sense than in a discipline-autonomous sense. According to him, the functioning of a discipline and its social and political environment determine the degree of autonomy of a scientific field. Autonomy is seen as the "ability to break external constraints or demands, to bring them into a specific form. [...] The decisive indication of the degree of autonomy of a field is, therefore, its refractive power, its power of translation." (Bourdieu, 1998, p.19, own translation). The social sciences, especially educational research, tend to incorporate the expectations of the different environments into their own terminology. As a "fractured-porous discipline" (Meusburger 2009, p. 117f.), it is characterized by weak disciplinary demarcation, which favors the import of theories and methods from other disciplines but has a high creativity and innovative potential precisely through this import. In this regard, however, there are subtle differences in disciplines as well as in

educational research cultures (Knaupp et al. 2014), and also in the selective reception, translation of knowledge, and subsequent adaptation to the national language and theory space (Hofbauer, 2018). German educational research has a high intensity of self-reflecting discourses on theoretical imports, disciplinary autonomy, and profile (as well in Austria, German-speaking Switzerland, and Belgium). These "*formalized disciplinary discourses*" contrast with the model of "*pragmatically specialized professions*" of educational research in Anglo-Saxon countries. These more pragmatic research cultures emphasize specific professional, political, and practical problems as well as economic, social, and political needs. As a result, research topics change dynamically and according to socio-political urgency, corresponding research groups emerge and produce knowledge which is also market-dependent through mediation (ibid.; Hofbauer, Kelly, & Beck, 2021). The models of research discourses are an expression of the social and intellectual character of educational research, which is shaped by particular national, cultural, and, in this context, linguistic backgrounds. Such discourse models prove to be relatively historically stable patterns of communication and interpretation, even though research cultures are characterized by mobility, knowledge circulation, competition, communication, and exchange with complex basic contexts on the one hand and synchronous and diachronic logics on the other (Werner & Zimmermann, 2006). Therefore, it can also be assumed that, despite inflationary and seemingly deliberate use of the term "evidence," misunderstandings and shifts of meanings occur especially in translations.

3.   Insights in German and English Educational Research Discourses

To show the interconnectedness of the use and establishment of the concept of evidence, we first trace the agenda and recommendations of European political actors. In the circulation of knowledge, science, through its international communication and exchange platforms (such as the *European Educational Research Association,* but also international politically figuring agencies, such as the *European Commission* and the *OECD*) have special importance.

The OECD is "a key supplier and interpreter of the type of evidence appreciated by politicians and decision-makers who can ascribe their narratives to" them (Ydesen, 2019, p. 2). In the following, we show OECD reports that use the term "evidence" to link educational research and the legitimization of educational policy decisions. However, it remained a broadly English term until the 1990s (3.1). The OECD, especially PISA, had little impact on policy-making in England. Rather, the debate on evidence-based education with comparisons to evidence-based medicine was triggered by an academic speech, which in turn was taken up by the OECD again (3.2). For the German-speaking discussion, it can be seen that the concept of evidence emerged particularly in the earlier 2000s, i.e., after the OECD reports of the 1970s and 90s (3.3). For the analysis, we explicitly looked for the term "evidence" and "*Evidenz*" in reports and texts, examining context, references, and usage. Iteration, reproduction and iterability, or the accompanying circulation, but

also the formation of tradition, thus come into the focus (Hofbauer 2023).

3.1   Better Research for Better Evidence: non-translations

Recommendations for improving education and educational research for a lot of member states have been issued by the OECD for some time. These recommendations are embedded in the attributed role of science and higher education as an instance of innovation in a knowledge society (OECD, 1972) on the one hand, and as an instance of growth and wealth creation on the other, along with human capital theory (OECD, 2001). In addition, the OECD discovered the indicator-based measurement and evaluation of science as early as the 1960s, using the US work of the *National Science Foundation* as a model (Godin, 2001). The evaluation of research must act in the contexts of efficiency, competition, productivity, political will for reform and technological progress with "value for money" (OECD, 1971; 1987). The reviews and recommendations became more discipline-specific over time. In particular, between 1995 and 2007, a series of expert meetings and reports emerged that revolved around the performances of research in education.[3] University based educational research (and development) of various countries were also evaluated. The OECD or representatives of empirical educational research made recommendations that called empiricism and evidence production to act as a counterbalance to the humanities-based "blue sky" science of education that was strongly represented at the university. Recommendations are therefore directed at many research forms,

---

[3] *Reviews of National Policies for Education* have been around since 1969, and the first reviews were on Sweden, Ireland, Italy, Austria, Holland, Japan, USA, and France. Germany was judges as being "deficient" in the education system in 1972 (OECD, 1972; 1973). Such reviews result from

one- or two-week visits, inspections, and discussions with scientific, educational, and school officials, which was then presented to the OECD Committee on Education. Since 1961, the focus has been on an economically oriented education policy (ibid).

although there is a need for more quantitative and qualitative evidence (Hofbauer 2020).

Although the OECD's evidence agenda has been widely discussed (Ydesen, 2019; Knejsler 2017), it seems that even in the 1990s the term was still stuck in the English-speaking tradition. The OECD report "*Bildung mangelhaft*" from 1973 announced that educational policy as well as research on education had received too little attention. "*Evidenz*" was not used once. The English term "evidence" was used in the 1995 OECD conference proceedings "Knowledge Bases for Education Policies" (p. 25-31), but was not used in the German volume published two years later ("*Wissensgrundlagen für die Bildungspolitik*," 1997, p.33-42). The author of the contribution was Donald Hirsch, an international policy consultant in the United Kingdom. He reported rather generally on the goals of the conference in discussion and the relation of educational research and development with further assumptions of the role of the *Centre for Educational Research and Innovation* (CERI). In the German translation, many more terms were used for the translation (indicators, results, information, findings, indications). In comparison, in the OECD documentation on the seminars about educational research and development in Austria, Germany, and Switzerland – also from 1995 – the term "evidence" was used but mainly from the report of the rapporteur Ivor Pritchard. Pritchard was a senior research analyst for the U.S. Department of Education at the time. So, even if the topic is the state of educational research and development in German-speaking countries, but in a report written in English, "evidence" only appears as a term if the author may have an English-language background.

In comparison, the debate on evidence-based teacher education and the role of research

erupted at the same time in England with consequences for the evaluative practices of the OECD in England itself and Denmark. England was under the scrutiny of the OECD in 2002, which was critically examined in 2003 by the *European Educational Research Association* (Wolter, Keiner, Paloma, & Lindblad, 2004). Like the other reports, it pleaded for more balanced research portfolio in promoting basic as well as useful research. Accordingly, "usefulness" has become a standard of quality (Keiner, 2005) and therefore affects and regulates research processes and discourses by providing a point of reference. The (English) OECD examination reports and recommendations for educational research are not only directed at its performance in terms of "better data," but they are also linked to the innovative capacity of the education system through the overarching demand for educational policy advice by evidence production. In comparison, Scotland's use of educational research has received less attention from the OECD. The most notable development here is the OECD's examination of the Scottish education system in 2015 and subsequent report, which argued strongly for the increased use of data and research evidence in education, along with an evidence-informed approach to policy (OECD, 2015, p. 12). In addition to this OECD science policy discourse, the agendas of the European Commission and the European Union between 2004 and 2020 also stand out, positioning evidence-based and education as guiding categories for the scientific justification of political decisions on the one hand and Europe as a knowledge-based, growth-oriented economic area on the other. Research-based evidence is placed alongside other forms of knowledge and assessment but ties the target categories to the coordinated expansion of infrastructural data generation in education and the evaluation of benchmarks. It is precisely this

triad of educational research, educational policy and education that constitutes an aspect of the international evidence movement that has far-reaching discursive, infrastructural, and epistemological consequences (Sellar & Lingard, 2014, Ydesen, 2019; Wilmers & Jornitz, 2018).

3.2   *Better Evidence for Better Teacher Education in England*

The discussion about "evidence"-informed teacher education has discursive antecedents in German-speaking as well as English-speaking educational research with regard to the question of knowledge- and research-based teaching and pedagogical action. In England, the market-based educational reforms that followed from 1988, mobilized through the introduction of teacher accountability measures linked to student tests in 1990 and regular visits from *Office for Standards in Education*, Children's Services and Skills (Ofsted), the school inspections agency, after 1992, afforded the emergence of a new technicist teacher professionalism. In 1996, David Hargreaves, comparing education and educational research to medicine, argued that educational research should attend more to gathering evidence of what works in what circumstances for use by teachers (Hargreaves, 1996a; 1996b). In his terms, evidence is specifically evidence of what works in improving student attainment. The evidence knowledge base for teachers should be more effective by allowing through diagnosis, the improvement of teaching and therefore become "satisfying" (1996a, ibid p. 1). Over time, as high stakes testing has become central to the evaluation of schools, this has become evidence of what works in improving high-stakes test results.

Hargreaves' call for an evidence-based teaching profession based on medicine (1996a; 1996b) allowed school improvement to become

synonymous with the disinterested pursuit of public wellbeing. The *Teacher Training Agency* (TTA), which played a huge part in reforming and reconceptualizing competences in teacher education as national standards (Reynolds, 1999) invited him to give the lecture. His speech on evidence-based teaching triggered a critical response and a debate (Hammersly 1997; Hargreaves, 1997; Elliott, 2001). Hargreaves's speech was so influential that it was discussed in the OECD's Evaluation of Educational Research and Development in England: "He compared educational research with medical research and found educational research deficient in important dimensions: it was non- cumulative, not useful for improving schools and generally lacking in quality" (OECD, 2002, p. 9; OECD, 2003, p. 26 & 63-64). Even more, he was also one of the interviewees whose statements were included in the evaluation (ibid. p. 55) and was appointed to a review team the educational research and development in Denmark (OECD, 2004, p.4).

Amidst a growing preoccupation with evidence-based and later evidence-informed practice, this included both understanding and making use of external research evidence to develop initiatives for tackling specific school issues, and internal evidence from student assessments and teaching observations to improve the focus and quality of teaching, something encouraged by Ofsted. This was clearly aligned with changes in the nature of teachers' professionalism in the context of "new public management," where the emphasis is on explicit measures of performance, output controls, market-type mechanisms and the introduction of competition (Bottery, 1996). During this time, several not-for-profit non-state intermediaries arose to make evidence of what works and examples of best practice available to school leaders and teachers to help them

improve their practice (Lawn & Furlong, 2007). These were established in parallel with an ongoing interest in school effectiveness research, which originated in the United States, and included the *Evidence for Policy and Practice Information Centre* (EPPI), founded in 1993, and, from 2008, the *Centre for the Use of Research and Evidence in Education* (CUREE). The school effectiveness movement (as opposed to school improvement which maintained a broader perspective) turned the focus of their research from understanding processes of teaching and learning to input-output evaluations of student outcomes. Beginning with David Reynolds in the 1980s (who later worked with Daniel Muijs, now head of research at Ofsted) and later including Peter Mortimore, they focused on identifying the characteristics of effective schools where pupils progress further than might be expected from consideration of its intake (Mortimore, 1991), mainly using large scale quantitative studies seeking statistical associations. In evaluating this movement, Chitty (1997) argues its weaknesses included failing to look at the broader policy context and placing too much emphasis on progressive school management as the dynamic of change, ignoring socio-economic status and offering no substantive detail on curriculum and pedagogy.

The Teaching and Learning Research Programme was established alongside these developments, and in direct response to Hargreaves call and the criticisms of other regulatory bodies of educational and pedagogical research as being small scale, irrelevant, inaccessible and of low quality (Parsons & Burkey, 2011). This was a £30 million initiative which ran from 1999-2009, managed by the Economic and Social Research Council, substantial in both scope and scale and including over 60 major projects involving more than 300 researchers. Starting as a broad

enterprise more interested in school improvement, in which researchers worked closely with practitioners to ensure the relevance and application of their findings to policy and practice and inform and enrich teaching and learning, in time it became more focused on research focused on increasing efficiency and effectiveness in pedagogy and in producing evidence informed pedagogic principles (James & Pollard, 2011).

In 2010, education secretary Michael Gove announced the coalition government had allocated £110 million to establish an education endowment fund to raise standards in underperforming schools. Renamed the *Education Endowment Foundation* (EEF), this has become the biggest funder of education research in the UK, and much of its funding goes to non-university-based researchers (Gorard, et al., 2017). The EEF, whose Teaching and Learning Toolkit has been available to schools since 2011, is currently one of the most influential evidence brokers in England and offers services in research, reviews, and consulting to provide "better" knowledge for "better" decisions. The EEF is a charitable organization which, according to its website, "seeks to break the link between family income and educational achievement" (EEF, 2020). It aims, amongst other things, "to raise the attainment of 3-18 year-olds, particularly those facing disadvantage by: (a) supporting teachers and senior leaders by providing free, independent and evidence-based resources – summaries and practical tools – designed to improve practice and boost learning; (b) generating evidence of what works to improve teaching and learning, funding rigorous trials of promising but untested programs and approaches; and (c) supporting schools, as well as early years and post-16 settings, across the country in scaling evidence to achieve the

maximum possible benefit for young people" (EEF, 2020).

Nordin and Sundberg (2014) show how international organizations like the OECD, which administers and analyses the PISA comparative tests, are now highly influential, as rising international governance and transnational activities have lessened state sovereignty. Within the OECD's *Centre for Educational Research and Innovation* (CERI) led by Tom Schuller, a British academic, they began to focus on evidence-based policy research from 2003 and, in 2006, hosted a series of conferences and meetings about evidence-based policy research in education. In these, they were heavily influenced by the TLRP who had begun working with researchers including Stephen Gorard to explore the use of randomized controlled trials in education (OECD, 2007). Throughout, the OECD's tone has been largely neoliberal (El Bouhali, 2015), increasing its role in recent years to provide explanations for national differences in PISA scores and to make policy recommendations (Sellar & Lingard, 2014).

Hargreaves's speech suddenly brought the concept of evidence into the consciousness of educational research production in the British debate. In doing so, it is embedded in years of reform and in the coupling of scientific, political, and practical sub-societal values by institutionalization of information centers and networks. As a result, politicians looking to gain public support for their policies, even those serving dominant interests, could select data or research to justify them, and thereby appear impartial and concerned for the welfare of everyone. Against this background, it also becomes clear how the question of the inner-scientific scope of scientific knowledge is less discussed than the effects that are of interest.

3.3   From Research Knowledge to "Evidence" in the German Educational Research Discourse

Against the background that even the OECD did not use the word "evidence" in the German context in the 1990s, this chapter presents an exploratory approach to the circulation of words in educational research. To get a better insight into the evidence circulation in German educational research, the central database of educational research was consulted. It shows how the use of the term "*Evidenz*" in German educational research has increased considerably in recent times (**figure 1**).

The steady increase in the use of "evidence" can be marked in the transition to the 2000s. Even if the respective context of term use cannot be discerned by such a simple count of a literature database. However, based on the 794 entries, at least thematizations can be read that are related to "evidence," which are contextualized with the (new) developments of the discipline and science policy funding to be considered.

The first publications recorded in the picture (1980-1990) concern psycholinguistics, language development and the history of poetry in the English classroom. Evidence is used as thesis-confirming, empirical knowledge. In 1993, the term then appears in a title on the interactions of social management and education in Austria. Also, from 1993 on, papers on cognitive psychological and neurobiological issues can be found. These neurological publications provide the link to social medicine and evidence-based medicine while keywords with "medicine" only appear from 2006 onward. In particular, the "secured" knowledge of neurological research and brain research, the implications for pedagogical processes with follow-up discussions of what human beings are,

what is morality and freedom, came to bear particularly at this time. Educational research in Germany as well as in English speaking countries has enthusiastically taken up and further differentiated neurological findings (Becker, 2006; Rose & Abi-Riched, 2014). [4] Until then, "evidence" came to belong to the communicative fringe of the discipline. From around 2000 onward, there is a focus on human services, quality management, education, and higher education from the perspective of new management, educational reform from an economic perspective (internal efficiency, competition, investment) but also on technological means such as the use of computers. After 2005, questioning, critical positions on evidence-based practice can also be found. It is interesting to note that the article on social work (Schnurr, 2005) speaks of the beginning, i.e., the *first debating* of an old theory-practice tradition with terms of treatment which were used in medicine. In parallel, however, promises of solutions to social inequality are also published that explicitly name "evidence" as a category of this solution. In the negotiation of both positions - i.e., contributions to evidence as an empirical, educational policy and practical solution and

evidence as a discussed irredeemable promise of salvation - it becomes clear how the term itself is used normatively as well as reflexively within educational research discourses. Thus, it moves into the consciousness and body of knowledge of the discipline.

Figure 1: *Search for "Evidenz" in the German Educational Research literature database FIS-Bildung (N=794)* [5]
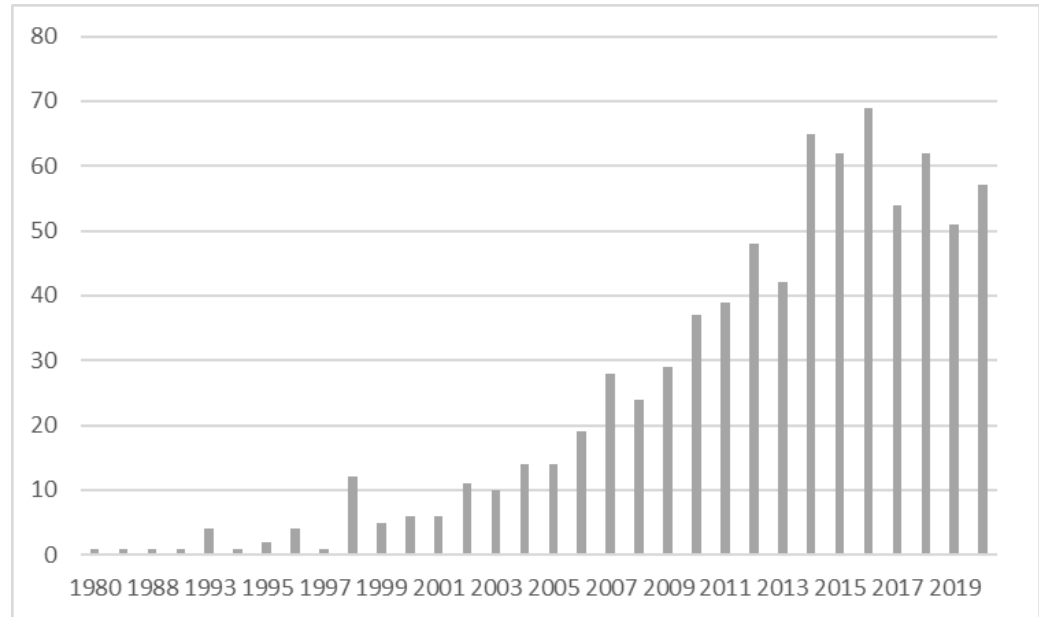
In addition, after 2000 the increase in educational, public, and scientific attention due to the PISA study can also be noted, at least in contemporary events (Ertl, 2006). Throughout the 1990s, a group of scholars who favored quantitative, empirical research methods labelled their type of research as *Empirische Bildungsforschung* (empirical educational research) with efforts to advance the "scientification" and "internationalization" with reference to global research standards (Keiner, 2015; Gross, Hofbauer, & Keiner, 2022). Large-scale assessments like PISA highlighted massive

---

[4] Similar for the reception in psychology, which, however, as a cognitive science was even more in competition with neuroscience, see Eckardt (2001).
[5] "*Evidenz*" is a keyword in *FIS-Bildung*. However, "evidence" was used as a general search term, so that insights into the titles and abstracts are also possible. *FIS-Bildung* also accesses English-language databases. The

results were limited by the specification on German publications. A comparison with English-language databases was not made, as the data volume of English-language publications (for example, "evidence" limited to "education" at Springerlink: 94724 entries; or ERIC: 88095 entries) would not be manageable.

gaps in pupils' test results, participation opportunities in education and social mobility. Through the accompanying massive media attention on the educational failure of an entire nation, the study pointed at quality issues of schools and teaching that were too little observed and too little observable in a standardized way. This meant that it was no longer just about the performance of the education system per se, but additionally about the performance of a science that is dedicated to improving the structures of educational processes. PISA therefore not only became a discourse and research catalyst for the empirical research traditions (Zapp & Powell, 2016), philosophical representatives of educational research also took up PISA references within the semantics on educational justice – even if these references were made through criticism and demarcation (Vogel, 2016). This is also reflected in the publications of the literature database. In 2002, the first publication with reference to PISA and "evidence" appears, which deals with the reading performance of German pupils. From then on, publications with reference to "evidence" increase - especially in the titles. Critical perspectives also increase, which highlight the lack of subject orientation and pedagogical goal perspective in educational research due to the strong test and data focus. Moreover, the scientific and public attention to PISA and the increased use of the concept of evidence is accompanied by funding for empirical educational research and processes of demarcation from science in education at the national level. Such movements trigger criticism and skepticism, so that representatives of empirical educational research saw themselves in the position of responding to this criticism, to highlight the performance and limitations of evidence production *within the scientific community* (Baumert & Tillmann, 2016). The limitations are found particularly in the

initiation of decisions and the uncertain implementation of decisions made outside the scientific scope. Thus, "evidence" remains for the time being only on the side of the scientific production community itself and not on the side of the "effect."

4.   The Circulation and (Mis)Understanding of "Evidence"

Although the concept of evidence emerged as early as the 17th Century as an epistemological and methodological concept, it is also as an expression of scientific efforts to improve knowledge production. Moreover, while the term circulated widely in and between the sciences, it seems not to appear in the (German) philosophically oriented, humanistic science on education (beside the critics). In this context, links to what is visible (*ersichtlich*), the visualisation through representation (*Darstellung*), could even be connected to a subject-oriented conception of *Bildung* (self-formation) as well as science. Rather, the concept of evidence marks a disciplinary development that is distinct from the non-methodological if one considers the emergence and self-image of empirical educational research. While in German-speaking countries PISA strengthened the incipient use of evidence after the reception of neurological expertise, the "need for more evidence" with "value for money" (Hargreaves, 1996a) discussions started in England before PISA, but influenced OECD recommendations on an international level.

Of particular interest in the OECD analysis was the insight that "evidence" was translated into various other German terms. The fact that the terms are used synonymously from English into German was apparently not common at all for a long time. Terms and concepts of knowledge also have their own historical, linguistically, and culturally

embedded "theory space." This seems less obvious because such terms convey a supposedly clear idea of scientificity. The direct translation "evidence" into "*Evidenz*" since the 90s and early 2000s without problematization illustrates this. Nevertheless, subtle differences in usage can be seen based on the respective academic traditions and how they justify and position the need for evidence. In German-speaking empirical educational research the critical examination of "evidence" is given priority. The "formalized disciplinary discourse" thus remains, despite the possibility of coupling scientific test, effect, and educational policy. It remains bound to internal disciplinary communication and justification. In comparison, the evidence movement in England, set within the model of "pragmatically specialized professions," has an institutional momentum of its own and is focused on effect in terms of what works and best practice. This means that although the term "evidence" has many theoretical, methodological, linguistic, social, and political possibilities for connections, the respective academic patterns of communication and interpretation stay relatively stable. Rather, despite an increased reference to a circulating notion of knowledge that conveys the idea of scientificity but also demonstration and effect, the potential for misunderstanding between educational researchers on an international level is not reduced even if the "evidence" – "*Evidenz*" translation suggests this.

References

Baumert, J. & K.-J. Tillmann (2016). *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker* (31. Sonderheft der Zeitschrift für Erziehungswissenschaft, pp. 129-146). Wiesbaden: Springer VS.

Becker, N. (2006). *Die neurowissenschaftliche Herausforderung der Pädagogik*. Bad Heilbrunn: Klinkhardt.

Biesta, G. (2010): Why 'What Works' Still Won't Work: From Evidence-Based Education to Value-Based Education. *Studies in Philosophy and Education, 29*, 491-503.

Bottery, M. (1996). The challenge to professionals from the New Public Management: implications for the teaching profession. *Oxford Review of Education, 22*, 179–197.

Bourdieu, P. (1998). *Vom Gebrauch der Wissenschaft. Für eine klinische Soziologie des wissenschaftlichen Feldes*. Konstanz: Universitätsverlag Konstanz.

Bourdieu, P. (2002a). Les conditions sociales de la circulation internationale des idées. *Actes de la recherche en sciences sociales, 145*(5), 3-8.

Bourdieu, P. (2002b). *Ein soziologischer Selbstversuch*. Frankfurt a. M.: Suhrkamp.

Campe, R. (2010). Shapes and Figures. Geometry and Rhetoric in the Age of Evidence, *Monatshefte, 102*(3), 285-299.

Campe, R. (2015). Epoche der Evidenz. Knoten in einem terminologischen Netzwerk zwischen Descartes und Kant. In S. Peters, & M. J. Schäfer (Eds.), *»Intellektuelle Anschauung«* (pp. 25-43). Bielefeld: transcript Verlag.

Chandler, J.; Davidson, A. I. & Harootunian, H. D. (1994). *Questions of evidence: proof, practice, and persuasion across the discipline*. Chicago: Univ. of Chicago Press.

Chitty, C. (1997) The school effectiveness movement: origins, shortcomings and future possibilities. *The Curriculum Journal, 8*(1), 45-62.

Colucci-Gray, L., Das, S., Gray, D., Robson, D. & Spratt, J. (2013). Evidence-based practice and teacher action-research: A reflection on the nature and direction of "change." *British Educational Research Journal, 39*. 126-147.

Cooper, A., & Levin, B., & Campbell, C. (2009). The Growing (But Still Limited) Importance of Evidence in Education Policy and Practice. *Journal of Educational Change. 10*. 159-171.

Eckard, G. (2001). Anmerkungen zur Relation von disziplinärer Perspektive und interdisziplinärer Verflechtung aus psychologiehistorischer Sicht. *Psychologie und Geschichte, 9*(3-4), 128-138.

EEF (2020). *Our mission*, https://educationendowmentfoundation. org.uk/about/history/ [Accessed 01 March 2022].

El Bouhali, C. (2015). The OECD Neoliberal Governance: Policies of International Testing and their Impact on Global Education Systems (pp. 119-129). A. A. Abdi, L.Shultz, & T. Pillay (Eds), *Decolonizing Global Citizenship Education*. Rotterdam: Sense.

Elliott, J. (2001). Making Evidence-Based Practice Educational. *British Educational Research Journal, 27*(5), 555–574.

Ertl, H. (2006). Educational standards and the changing discourse on education. The reception and consequences of the PISA study in Germany. *Oxford Review of Education, 32*(5), 619-634.

Fleck, L. (1979). Genesis and Development of a Scientific Fact. Chicago: The University of Chicago Press.

Fuller, K., & Stevenson, H. (2019). Global education reform: understanding the movement. *Educational Review, 71*(1), 1-4.

Godin, B. (2001). The emergence of science and technology indicators: why did governments supplement statistics with indicators? *Project on the History and Sociology of S&T Statistics*, Paper No. 8. Online: http://csiic.ca/PDF/Godin_8.pdf. [Accessed 27th March 2022].

Gorard, S., See, B.H., & Siddiqui, N. (2017). *The trials of evidence-based education: The promises, opportunities and problems of trials in education*. London: Routledge.

Gough D. (2021). Appraising Evidence Claims. *Review of Research in Education. 45*(1),1-26.

Grey, S. (2020). *PISA, the media and the governance of education policy in England and Sweden*. Doctoral Thesis. UCL.

Gross, B., Hofbauer, S., & Keiner, E. (2022). The "Science of Education" – Different Terms, Concepts, Cultures and Epistemologies? A Contribution to a Social Epistemology. *Società di Politica Educazione e Storia Rivista di Politica, Educazione e Storia. N. 16, 19-38.*

Hammersly, M. (1997). Educational research and teaching: a response to David Hargreaves's TTA Lecture. *British Educational Research Journal, 23*, 141-161.

Hargreaves, D. (1996a). Educational research and evidence-based educational practice: A response to critics. *Research Intelligence, 58*, 12–16.

Hargreaves, D. (1996b) *Teaching as a research-based profession: possibilities and prospects. Teacher Training Agency Annual Lecture*, London, Available at: https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/TTA%20Hargreaves%20lecture.pdf [Accessed 27th March 2022].

Hargreaves, D. (1997). In defence of research for evidence-based teaching: a rejoinder to Martyn Hammersley. *British Educational Research Journal, 23*. 405-419.

Helgetun, J .B., & Menter, I. (2022). From an age of measurement to an evidence era? Policy-making in teacher education in England, *Journal of Education Policy, 37*(1), 88-105.

Hofbauer, S. (est. 2023): Vom Wort zur Wiederholung – methodologische und methodische Zugänge in der Analyse erziehungswissenschaftlicher Sprache. In: A. Böhmer; U. Binder & J. Oelkers (eds.): Sprache und Pädagogik.

Hofbauer, S. (2020): Von *Action Research* zu *Policy Experimentation in Education*. Die Verschränkung von Bildungsforschung und Bildungspolitik in Dokumenten der OECD/ CERI seit den 1990er Jahren (pp. 157-170). In U. Binder, & W. Meseth (eds), *Strukturwandel in der Erziehungswissenschaft. Theoretische Perspektiven und Befunde*. Bad Heilbrunn: Klinkhardt.

Hofbauer, S. (2018): Konstruktionswandel und kollektive Übersetzungsmacht. Erziehungswissenschaftliche Formierung des Forschungsgegenstandes „Lehrerwissen" in diachroner Perspektive (pp. 107-126). In M. Lüders, & W. Meseth (eds.), *Theorieentwicklungen in der Erziehungswissenschaft. Befunde – Problemanzeigen – Desiderata*. Bad Heilbrunn: Klinkhardt.

Hofbauer, S.; Kelly, P, & Beck, A. (2021). Between Inclusion and Improvement: Evidence and the Politics of Teaching Assistants in England (pp. 213-225). In D. Kemethofer, J. Reitinger & K. Soukup-Altrichter (eds), *Vermessen?! Zum Verhältnis von Bildungsforschung, Bildungspolitik und Bildungspraxis*. Münster: Waxmann.

James, M., & Pollard, A. (2011). TLRP's ten principles for effective pedagogy: rationale, development, evidence, argument and impact. *Research Papers in Education, 26*(3), 275-328.

Jurt, J. (2014). Die sozialen Bedingungen der internationalen Zirkulation der Ideen. Das Konzept Pierre Bourdieus und dessen Applikation (pp. 155-169). In S. Schmidt (eds), *Wissen in Bewegung: Theoriebildung unter dem Fokus von Entgrenzung und Grenzziehung*. Berlin: Lit Verlag.

Keiner, E. (2005). Zur Konstruktion erziehungswissenschaftlicher Forschung aus der Perspektive der OECD. In I. Gogolin I.; HH. Krüger; D. Lenzen, & T. Rauschenbach (eds.), *Standards und Standardisierungen in der Erziehungswissenschaft. Zeitschrift für Erziehungswissenschaft* (4. Beiheft Zeitschrift für Erziehungswissenschaft, pp. 13-22), Wiesbaden: VS Verlag.

Keiner, E. (2015). Pädagogik, Erziehungswissenschaft, Bildungswissenschaft, Empirische Bildungsforschung – Begriffe und funktionale Kontexte. In E. Glaser, & E. Keiner (Eds.), *Unscharfe Grenzen – eine*

Disziplin im Dialog (pp.13-34). *Pädagogik, Erziehungswissenschaft, Bildungswissenschaft, Empirische Bildungsforschung*. Bad Heilbrunn: Klinkhardt.

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

Krejsler J.B. (2017). Capturing the "Evidence" and "What Works" Agenda in Education: A Truth Regime and the Art of Manoeuvring Floating Signifiers. In M. Eryaman, & B. Schneider (eds), *Evidence and Public Good in Educational Policy, Research and Practice. Educational Governance Research*, vol 6. Cham: Springer.

Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lawn, M., & Furlong, J. (2007). The Social Organisation of Education Research in England. *European Educational Research Journal*, *6*(1), 55-70.

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.

Meusburger, P. (2009). Milieus of creativity: The role of placed, environments, and spatial contexts. In P. Meusburger, J. Funke & E. Wunder (Eds.), *Milieus of creativity: An interdisciplinary approach to spatiality of creativity* (pp. 97-154). Dordrecht: Springer.

Mortimore, P. (1991) School effectiveness research: which way at the crossroads? *School Effectiveness and School Improvement, 2*(3), 213-229.

Nordin, A., & Sundberg, D. (2014) (Eds). *Transnational policy flows in European education: the making and governing of knowledge in the educational policy field*. Oxford: Symposium Books.

OECD (1987): *Evaluation of current research. A selection of current practices*. Paris: OECD.

OECD (1972). *Germany. Reviews of National Policies for Education*. Paris: OECD Publishing.

OECD (1971). *Science Growth and Society: A New Perspective*. Paris: OECD Publishing

OECD (1973). *Bildungswesen mangelhaft. BRD-Bildungspolitik im OECD-Länderexamen*. Frankfurt am Main: Diesterweg.

OECD (1997). *Wissensgrundlagen für die Bildungspolitik*. Paris: OECD Publishing.

OECD (2001). *The wellbeing of nations. The Role of Human and Social Capital*. Paris: OECD Publishing.

OECD (2002). *Educational Research and Development in England. Examiners' Report on England*. Paris: OECD Publishing.

OECD (2003): *New Challenges for Examiners' Report on England. OECD Publishing Educational Research*. Paris: OECD Publishing.

OECD (2004). *National Review on Educational Research and Development. Examiners' Report on Denmark*. Paris: OECD Publishing.

OECD (2007) *Evidence in Education: Linking Research and Policy. Paris:* OECD Publishing.

OECD (2015). *Improving Schools in Scotland: An OECD Perspective*. Paris: OECD Publishing.

Parsons, D., & Burkey, S. (2011). *Evaluation of the Teaching and Learning Research Programme (Second Phase): Final Report*. Horsham: HOST Policy Research.

Pollitt, C., & Bouckaert, G. (2011). *Public management reform, a comparative analysis: New public management, governance, and the neo-Weberian state* (Third Edition), Oxford: Oxford University Press.

Reynolds, M. (1999). Standards and Professional Practice: The TTA and Initial Teacher Training. *British Journal of Educational Studies, 47*(3), 247–260.

Rose, N. & Abi-Rached, J. (2014). Governing through the Brain. Neuropolitics, Neuroscience and Subjectivity. *Cambridge Anthropology, 32*(1), 3–23.

Sarasin, P. (2011). Was ist Wissensgeschichte? *Internationales Archiv für Sozialgeschichte der deutschen Literatur, 36*(1), 159-172.

Sellar, S. & Lingard, B. (2014). The OECD and the expansion of PISA: new global modes of governance in education. *British Educational Research Journal, 40*, 917-936.

Schnurr, S. (2005): Evidenz ohne Reflexivität? - Zur Debatte um Evidenzbasierte Praxis in der Sozialen Arbeit. *Forschung & Wissenschaft Soziale Arbeit, 2*, 19-28.

Tseng V., & Nutley S. (2014) Building the Infrastructure to Improve the Use and Usefulness of Research in Education. In K. Finnigan & A. Daly (eds) *Using*

Research Evidence in Education. Policy Implications of Research in Education, vol 2. Cham: Springer.

Vogel, K. (2016). *Konstruktionen und Rezeptionen erziehungswissenschaftlichen Wissens. Bibliometrische und systematische Analysen am Beispiel des Diskurses ‚Bildungsgerechtigkeit'*. Bad Heilbrunn: Verlag Julius Klinkhardt.

Werner, M., & Zimmermann, B. (2006). Beyond Comparison: Histoire Croisée and the Challenge of Reflexivity. *History and Theory*, 45(1), 30-50.

Wilmers, A., & Jornitz, S. (2018). Transfer zwischen Bildungsforschung und Bildungspolitik: die europäische Dimension und Diskussion. In K. Drossel &; B. Eickelmann (eds.), *Does "What works" work? Bildungspolitik, Bildungsadministration und Bildungsforschung im Dialog* (p. 191-205). Münster: Waxmann.

Wolter, S., Keiner, E., Palomba, D., & Lindblad, S. (2004). OECD Examiners' Report on Educational Research and Development in England. *European Educational Research Journal, 3*(2), 510-526.

Ydesen, C. (2019). Introduction: What Can We Learn About Global Education from Historical and Global Policy Studies of the OECD? (p. 1-14) In ibid. (eds.), *The OECD's Historical Rise in Education. The Formation of a Global Governing Complex*. Palgrave.

Zapp, M., & Powell, J. J. W. (2016). How to construct an organizational field: Empirical educational research in Germany, 1995–2015. European Educational Research Journal, 15(5), 537–557. https://doi.org/10.1177/1474904116641422