

# Moving towards Mode 2? Evidence-based policy-making and the changing conditions for educational research in Germany

Mike Zapp\* and Justin J. W. Powell

Institute of Education and Society, University of Luxembourg, Luxembourg

\*Corresponding author. Email: mike.zapp@uni.lu

## Abstract

The ‘Mode 2’ approach is among the most widely used to analyze changes in contemporary science and innovation systems. This approach suggests that application-driven, transdisciplinary, reflexive, and contextualized scientific knowledge will be produced by an increasingly heterogeneous set of organizations, with universities no longer as dominant. Analyzing the case of educational research (ER) in Germany, which has undergone profound institutional and paradigmatic change since 2000, allows us to ask whether the Mode 2 thesis holds. Considerable investments in ‘empirical’ research and the top-down setting of the research agenda have, we argue, fundamentally altered the research infrastructure of this increasingly diverse multidisciplinary field, challenging the traditional humanities-based *Pädagogik*. Especially based on waves of large-scale assessments of school performance, the rapidly-growing ‘empirical’ ER field is characterized by quantitative and policy-relevant (applied) knowledge claims. Finally, we identify risks associated with rapid and policy-induced shifts in ER from Mode 1 to Mode 2.

**Key words:** educational research; Germany; policy-induced knowledge production; Mode 2; evidence-based policy-making.

## 1. Introduction

Over the past two decades, research and innovation systems have repeatedly been described as undergoing fundamental changes in their rationales, missions, organization, and governance. Concepts such as ‘academic capitalism’ (Slaughter and Leslie 1997; Slaughter and Rhoades 2009), the ‘triple helix’ of university–industry–government relationships (Etzkowitz and Leydesdorff 1998; Leydesdorff and Meyer 2006), or ‘post-academic science’ (Ziman 2000) depict various causes and consequences of such changes. These frameworks envision different roles for the university, perhaps the key organizational form of the ‘schooling society’ (Baker 2014). One of the most widely discussed proposals is the ‘new production of knowledge’ or ‘Mode 2 science’ (Gibbons et al. 1994; Nowotny et al. 2001). This framework assumes a shift from an academic, disciplinary, and autonomous university-based organization of primarily fundamental knowledge—described as Mode 1—to a more diverse, transdisciplinary, applied, and reflexive kind (Mode 2). Yet empirical cases, of certain fields in particular contexts, are needed to test the framework’s broader applicability. Acknowledged as an important conceptual approach to assess transformations in the organization of research, Mode 2 has also been critiqued for lacking empirical validity, among other concerns (see, e.g., Hicks and Katz (1996); Weingart (1997); Shinn (1999); Crompton (2007)). Here,

we address this lacuna by operationalizing Mode 2 concepts in the context of an empirical case we argue is particularly germane. Applying key tenets of the Mode 2 conceptualization to the field of educational research (ER) in Germany from 1995 to 2015, we analyze the changing conditions for ER, such as evidence-based policy-making, and uncover the extent to which these are adequately captured by Mode 2 tenets.

Increasingly, education has received considerable policy attention around the world. This interest has not only changed educational *systems* and their governance, but also transformed the perceived goals and functions of ER and the *kinds* of research to be promoted. Prominently, international large-scale assessments have grown in significance worldwide as aggregate educational performance is understood as central to reach a host of social, political, and economic goals (Heyneman and Lykins 2008; Benavot and Meyer 2013). Simultaneously, the notion of evidence-based policy-making has spread quickly around the world (Banks 2009; Dederig 2009; Howlett 2009). Evidence-based policy-making (EPM) in education is strongly anchored in specific kinds of ER. Often, this research is called empirical, quantitative, and interdisciplinary or multidisciplinary. Much of this research is reasoned causally and oriented to problem-solving. This description of ER may not seem particularly novel since many, especially Anglophone, countries have developed this kind of empirical educational research (EER) for decades. Yet

ER communities in other countries build on very different traditions, with Germany representing the paragon of a hermeneutics-trained and humanities-based pedagogy or educational scholarship with origins in 19th century philosophy. Yet since the turn of the century, a remarkable ER infrastructure has emerged there (Zapp and Powell 2016). This new organizational field, in contrast to the once dominant humanities-oriented pedagogy focused on *Bildung* (Hamann 2015; Horlacher 2016), is instead devoted to EER based mainly in the (quantitative) social sciences and psychology (see Terhard 2016).

On the basis of this selected case, we argue that this new research field displays key features that indeed characterize ‘Mode 2’ science. In the following, we introduce the conceptual underpinnings of Mode 2 science and present the analytical framework used to examine the changes in the organization of ER in Germany (Section 2). Section 3 discusses the methodology of the analysis and quantitative and qualitative data sources. Then, we present the findings (Section 4). In Section 5, we discuss several risks associated with such disruptive and policy-induced change to ER that have received little attention in the literature thus far. These threats concern changes in research management, the implications of the rising importance of alternative quality criteria, and the long-term sustainability of large-scale research programs initiated under Mode 2 conditions. We conclude with an outlook on future analytical strategies in uncovering changes in knowledge production (Section 6).

## 2. Characteristics of ‘Mode 2’ knowledge production

Soon after its publication, the *New Production of Knowledge* (NPK; Gibbons et al. (1994)) became an influential concept to describe the changing structures, conditions, practices, and outcomes of scientific work. In such fields as science and technology or innovation studies, the number of articles referring to the NPK doubled every two years by the mid-2000s (Hessels and van Lente 2008). Such attention is due to the prominence of the authors, surely, but also because of the universality of the claims therein (applying to the entire scientific system) and the disputability of their arguments. The main distinctions between Mode 1 and Mode 2, as outlined in Gibbons et al. (1994), precedes discussion of their more recent contributions.

### 2.1 Context of application

Knowledge production under Mode 1 conditions is characterized as ‘pure’, ‘theoretical’, ‘experimental’, or exclusively ‘academic’, while Mode 2 knowledge is regarded as ‘applied’. Application implies a stronger emphasis on knowledge management, transfer, dissemination, and use. More precisely, application can occur through various contextualizations of research—weak (e.g. particle physics), middle range (e.g. Human Genome Mapping Project), and strong (e.g. muscular dystrophy) (Nowotny et al. 2001; Gibbons and Nowotny 2003). The alleged tendency to orient research toward applicable ends—its ‘finalization’—overlaps with other authors’ judgment about strong governance and public policy objectives in research resource allocation as a key challenge in public science systems (van den Daele et al. 1979; Elzinga 1985, 1997; Whitley 2011). Has ER in Germany remained pure or become more applied?

### 2.2 Multi-/Transdisciplinarity

In contrast to Mode 1, knowledge in Mode 2 is produced trans- or multidisciplinarily. This entails the ‘mobilization of a range of theoretical perspectives and practical methodologies to solve problems’

(Gibbons and Nowotny 2003: 186). Transdisciplinary research requires more than just the collaboration of various disciplines (multi- or pluridisciplinarity) or the consolidation of overlapping research strands (interdisciplinarity). Instead, what distinguishes transdisciplinarity from its siblings is that it requires neither prior disciplinary knowledge order nor necessarily leads to a new order. Its external (social, political, economic) stimulus and internal (scientific) reactivity matter more than academic connectivity. However, the authors admit that transdisciplinarity is not easily achieved and that multidisciplinarity is its analytic predecessor, applied here as an additional heuristic strategy. Has German ER grown beyond one key discipline?

### 2.3 Organizational diversity

Sites of knowledge production under Mode 2 conditions multiply and interactions in expanding networks intensify. Beyond universities, governments, industry, international organizations (governmental and nongovernmental alike), think tanks, consultancies, associations, and activist groups join the knowledge-producing enterprise. Such diversity of organizational forms engaged in knowledge production contrasts with the traditional academic dominance of the university in Mode 1. Which organizational forms produce most ER in Germany?

### 2.4 Reflexivity and accountability

The new production of knowledge is said to be carried out dialogically between research actors and research subjects. Accountability stretches well into societies, beyond the immediate context of application. Investigations of natural and social worlds are now more recursively than objectively done (Mode 1). Sociological theories of ‘reflexive modernity’ (Beck et al. 1994) are echoed in later work (Nowotny et al. 2001). Scientific research is said to be subjected to new questions about its wider relevance and contributions, leading to new forms of institutional and individual reflexivity of scientific work freed from pre-defined academic codes. Has ER in Germany become more accountable and reflexive?

### 2.5 Novel quality control

Sources and forms of quality control multiply in Mode 2. Due to transdisciplinary complexity and wider social implications, scientific peers are no longer the only valid reference for quality control (as in Mode 1). Other knowledge producers and users are taken into account, including governments and society itself. Knowledge production is to become ‘socially robust’ or ‘context-sensitive’ (Nowotny et al. 2001: 117). How is quality guaranteed in contemporary ER in Germany?

## 3. Applying the Mode 2 concept to educational research

While many of these dimensions of the Mode 2 concept address important trends and patterns in knowledge production and science productivity, it is necessary to empirically analyze and reconstruct the shift from the traditional Mode 1 type to Mode 2. Indeed, of the various critiques Mode 2 has garnered, the concept’s lack of broad empirical validation is perhaps the most important (see, e.g., Hicks and Katz (1996); Weingart (1997); Shinn (1999); Crompton (2007)). The often-suggested historical order—that Mode 2 is contemporary, whereas Mode 1 is science’s original form—has been contested. Instead, some regard Mode 2 as the original conditions of

science, while Mode 1 is regarded as a political construct of the late 19th and early 20th century (Etzkowitz and Leydesdorff 2000; Rip 2000; Pestre 2003). The current transformation would, thus, mean a restoration of a previously distorted scientific order.

Other authors point to the need to further specify national and scientific disciplinary settings to do justice to the varied dynamics, challenging the universality of the proposed shift (Hicks and Katz 1996; Weingart 1997; Albert 2003; Tuunainen 2005). The lack of empirical evidence supporting the claim of the general rise of Mode 2 applies to the social sciences in particular. Godin (1998: 472), for example, argues that the 'social sciences, as well as the humanities, have always been of Mode 2, much more than has been the case for the natural and physical sciences'. He particularly contests the distinction between fundamental and applied research, pointing to the permanent mutual enrichment of these two forms of knowledge production.

Mode 2 proponents have, on the whole, not fully prepared Mode 2 to be analytically and methodologically applied, with the vast majority of works referring to Mode 2 simply using its tenets to frame arguments; less research applies these dimensions as analytical tools (Hessels and van Lente 2008). Among these, contributions to science, technology, and innovation studies dominate as well as do those focusing only on individual aspects of the new production of knowledge (e.g. transdisciplinarity, see Hicks and Katz (1996)). Applications within the humanities or social sciences (where the concept emanates from) are largely missing, although Gibbons et al. (1994) deliberately claim that all disciplines are equally affected by these general shifts. An exception is Albert's (2003) study on Canadian sociology and economics departments. The interviewed researchers and publications analysis do not yield substantial support for a Mode 2 shift, but rather the continuing importance of Mode 1.

Another Mode 2 application to the social sciences is Kropp and Blok's (2011) account of the long-term changes within Danish sociology, charting a growing resemblance with the Mode 2 form of knowledge production. They find substantial evidence for a Mode 2 shift of Danish sociology in most of the conceptual dimensions, except increasing transdisciplinarity. They depict the transformation in Danish sociology as incremental, largely mirroring changing 'priorities of the welfare state, in terms of politically prescribing and legitimating particular social role(s) of sociologists, social scientists, and science in general' (Kropp and Blok 2011: 223). The authors conclude by calling for a 'more historically grounded, empirically sensitive, and conceptually refined approach to studying these changing science–society relations' (Kropp and Blok 2011: 223).

Thus, here we attempt to provide such analytical thrust by applying the concept to an empirical case not yet well-researched in discussions about Mode 2, namely ER, whose origins in Mode 1 we briefly describe in the next section.

#### 4. Educational scholarship in Germany in Mode 1

First, the scientific community in Germany devoted to studies in education represented deeply-rooted scholarship of educational theory, philosophy, and history with different (sub-)disciplines such as pedagogy, didactics, and educational science. This highly idiosyncratic amalgam followed a markedly different approach from those evolving elsewhere (Biesta 2015). For six decades, Germany's education scholars and researchers have been organized mainly within one large disciplinary organization, the German Educational Research Association (DGfE). Its publication outlet was strongly traditionalist and insular. Its empirical section was complemented with members

from sociology and psychology, among other disciplines. This section, which would become so important from the mid-1990s onwards, had long remained marginal (Tröhler 2014).

Second, not only was its disciplinary scope hermetic, but also its scholarly communications and collaborations, such as publications and research networks, rarely crossed national or linguistic borders, that is, beyond the German-speaking world. The preferred type of publication was sole-authored monographs. International standards and processes in peer-review were largely nonexistent in academic journals dealing with education (Aljets 2015). Nor did Germany's ER community show much interest in joining international debates on achievement testing and performance assessment; now routine practice. Germany hardly participated in large-scale assessments until the mid-1990s, thus reducing opportunities for quantitative researchers to contribute analyses (Aljets 2015; Zapp and Powell 2016).

Third, awkward political entanglements under Nazi rule led to educational scholarship quite wary of close linkages to policy-makers, not only in the immediate aftermath of World War II. This view solidified, moving ER further away from policy-making circles in the 1960s, which first saw the revival of humanities-based pedagogy of pre-Nazi Germany and thereafter, the rise of critical education science (Biesta 2015; Horlacher 2016). Both strands strongly opposed understandings of education based on experts and engineering (both in its cognitive-psychological and behavioral variants) as established in many other Western countries and international organizations (Tröhler 2014).

The number of education professors was and still is comparatively high in Germany. While independent research institutes were scarce, they were influential and, initially, pursued research in opposition to the DGfE thrust. Many key research institutes today were founded not long after World War II: The German Institute for International Educational Research (*Deutsches Institut für Internationale Pädagogische Forschung*, DIPF, 1951), the Max Planck Institute for Human Development and Education (*Max-Planck-Institut für Bildungsforschung*, MPIB, 1963), and the Institute for Science and Mathematics Education (*Institut für die Pädagogik der Naturwissenschaften und Mathematik*, IPN, 1966), to name a few. They were and remain 'fact-based research bodies', joining wider calls for a 'realistic turn in pedagogical research' (Roth 1963: 109) that provided important empirical scholarship alongside certain universities such as Konstanz and Hamburg. Yet these institutes were also led by historical, comparative, or philosophy-oriented directors, reflecting the wider research tradition in German educational scholarship organized within the well-established discipline of *Pädagogik* (Ingenkamp 1992; Keiner 1999; Tröhler 2014).

Then, in the late 1990s, a growing number of science policy actors, such as the Science Council (*Wissenschaftsrat*, WR), the German Research Foundation (*Deutsche Forschungsgemeinschaft*, DFG), and the Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*, BMBF), called for a new model of ER based on an empirical-analytical, mostly quantitative 'exact social science' (WR 2001: 33; also DFG 2002). Consequently, humanities-based education scholarship was regarded not as the key contributor to this new research field, but rather as a hindrance. The BMBF, traditionally seen as a soft governance actor or coordinator in education in a highly fragmented federal system, has become surprisingly active in promoting (a certain kind of) ER (Edler and Kuhlmann 2008). It is convinced that 'empirical educational research is markedly different from the conventional work done in the more humanities-informed (school) pedagogy' (BMBF 2008: 8).

With this preliminary evidence, there is much reason to believe that Mode 2 conditions have the potential to alter the production of knowledge in *social scientific* fields as new social, political and economic demands grow (Kropp and Blok 2011). We argue that German ER constitutes an exemplary case to illustrate such changing conditions and consider the NPK hypothesis particularly apt to capture it, as it unites several analytical dimensions (e.g. content and governance) and levels (e.g. individuals, organizations, policy-making, and society).

The observation period under study (2000–15) includes the so-called PISA ‘shock’ since 2000–1 that has transformed German education (Waldow 2009): Mediocre results of German students in the OECD’s Programme for International Student Assessment (PISA) led to policy-makers initiating multiple reforms simultaneously. German ER was similarly transformed, providing a rich case to trace processes of institutional change in a predominantly public system of higher education. The changes were profound, with ER in the period prior markedly different from its contemporary shape. Fifteen years provide a sufficiently long span of time to trace the field’s change dynamics, involving various policy actors’ influence and research organizations’ reactions. Educational research showcases public intervention in science with education gaining signal importance in policy-making. Thus, we can trace mechanisms of direct government intervention and begin to estimate the consequences of such policy-induced Mode 2 transition.

## 5. Data, methods, and analytic strategy

Our mixed methods analysis draws on various sources of data, iteratively integrated to reconstruct the shifting shape of ER. First, we use longitudinal data on the volume, source, distribution of, and rationale for ER funding in the period 2000–13 to chart *long-term trends* in funding priorities.

The second data source is semi-structured expert interviews (Bogner et al. 2009), reconstructing shifts in German ER from *insiders’ viewpoints*. We conducted ten interviews of approximately 90 minutes each with experts from educational science and research policy-making, with sampling facilitated by the prominence of key figures involved in field development over the period. Among policy-makers, we interviewed actors from major funding organizations and educational administrations. We also spoke with boundary-spanning actors occupying the crucial middle ground between research and policy-making in both national and international contexts, translating and adapting scientific findings and political priorities. We chose interviewees whose role in reshaping ER was significant throughout the observation period. Many of these actors have (had) multiple functions, such as researcher, advisory board member, project leader, and consultant. This ensured that our analysis encompassed several perspectives; however, to assure anonymity, organizational affiliations are omitted in the presentation and interviewees are coded by A for *actor* and a consecutive number.

Third, we draw on key research policy documents, such as research programs funded by the Ministry of Education (BMBF) and policy documents from organizational actors involved ( $N = 15$ ) to reconstruct the *national discourse* on ER in the period 2000–15. These include major statements from the WR, the DFG, state ministries of education and speeches by Ministers of Education, organizational self-portrayals of the key ER institutes in Germany and professional associations as well as lists of conference participants to identify national and international networks in the field of ER. All

interviews and documents cited have been translated by the authors from the original German.

Finally, we draw on quantitative and qualitative content analyses of a comprehensive set of ER projects ( $N = 150$ ) funded by the DFG and the BMBF in the period 2007–15, which sheds light on the *contents* of German ER. The starting point for project sampling is defined by the launch of the BMBF Funding Priority Higher Education Research and the Framework Program for the Promotion of EER, of which we analyzed the fifty most expensive projects. The DFG sample was also selected based on the funding volume (the top 50 ER projects from the period).

To the best of our knowledge, empirical work that explicitly defines Mode 2 tenets as researchable questions or testable hypotheses for our case of ER is lacking. Yet, if we want to test the empirical validity, historical novelty, and generality of this approach and, more importantly, to harness the heuristic potential of Mode 2 concepts, we need to translate these into analytical categories. Thus, we ask whether ER in Germany displays the hypothesized shift from Mode 1, university-based knowledge production to a more application-oriented, multidisciplinary, diverse, reflexive Mode 2 form subject to more than purely scientific quality criteria and that is carried out in diverse organizational forms.

We deductively derived overarching categories in advance, representing the NPK’s five main tenets (discussed in Section 2): context of application, multi/transdisciplinarity, organizational diversity, reflexivity and accountability, and novel quality control. Then, we used specific research questions to operationalize each main category. Finally, further categories or concepts and their specific indicators found in the materials were inductively gathered and continuously refined. Table 1 provides an overview of refined operationalizations applied to analysis of both quantitative and qualitative data.

We may not expect a radical transformation of ER scholarship or wholesale incorporation of all Mode 2 tenets, as both academic cognitive and organizational structures usually change gradually. Instead, we chart incremental and discontinuous change in the general ER landscape. Different strands of traditional and novel ER scholarship continue to coexist and even coalesce, as strands of ER develop at different speeds. While some features predicted by Mode 2 tenets may display clear contours, others cannot (yet) have reached tangible form or even show reverse dynamics. Crucially, Mode 2 features overlap as highly interdependent. For example, the context of application is embedded in the social contexts of reflexivity and accountability and often entails novel forms of quality control (Gibbons et al. 1994). Organizational diversity could spur transdisciplinarity as different organizational cultures complement each other, inter/nationally. The analytical categories alongside their operationalizations do not represent an exhaustive research strategy in the framework of Mode 2. Given the fact that Mode 2 has yet to be sufficiently transformed into a methodologically-amenable framework, our modest aim here is to provide preliminary evidence for the concept’s analytical potential. Our goal is less to falsify Mode 2 hypotheses than to harness its heuristic potential and to provide evidence of the concept’s analytical potential in a specific empirical case, to which we now turn.

## 6. The changing conditions for educational research in Germany

### 6.1 Context of application

Applicability of social scientific research is closely related to the idea of EPM, diffusing from health policy to other social and educational



**Table 1.** Mode 2 analytical framework applied to educational research.

Mode 2 tenet/analytical category	Research questions (working hypothesis)	Indicators; data sources
Context of application	<ul style="list-style-type: none"> <li>To what extent do educational researchers perceive themselves as serving society and/or policy-making?</li> <li>Do policy actors relate to ER as a 'service' sector?</li> <li>In which contexts is ER applied?</li> <li>What infrastructures and mechanisms are used to facilitate knowledge transfer and policy advice?</li> </ul>	<ul style="list-style-type: none"> <li>References to evidence-based policy-making in policy documents and statements.</li> <li>References to knowledge application and transfer in interviews.</li> <li>'applied' research rationales in funding criteria.</li> </ul>
Transdisciplinarity	<ul style="list-style-type: none"> <li>Do we find evidence for common methodological understanding, shared diffusion channels and the transfer of knowledge to policy-makers, administration, and practitioners?</li> <li>How do professional associations representing ER in Germany portray themselves in terms of their commitment to transdisciplinarity?</li> </ul>	<ul style="list-style-type: none"> <li>Organizational self-portrayals.</li> <li>Perception of organizational leaders.</li> <li>Disciplinary composition in ER projects.</li> </ul>
Organizational diversity	<ul style="list-style-type: none"> <li>In what does the ER infrastructure consist?</li> <li>In which types of organizations is ER conducted?</li> <li>How have organizational forms changed, gaining or losing in importance?</li> </ul>	<ul style="list-style-type: none"> <li>Founding of new organizations for research.</li> <li>Expansion of extant organizations in terms of staff, professorships, funding.</li> <li>Intra-/extramural connections and (inter)national networks in ER.</li> </ul>
Reflexivity and accountability	<ul style="list-style-type: none"> <li>To what extent is research organized inter/nationally?</li> <li>What contexts (social, political, academic) shape ER?</li> <li>Do funding agencies require new forms of accountability?</li> <li>Which audiences take interest in ER?</li> <li>Are the researchers aware of these new stakeholders?</li> <li>Do scientific actors take on new roles as socially accountable servants and publicly discuss their research (strategies)?</li> <li>Have mass media become a forum for public debate and dissemination of research findings?</li> </ul>	<ul style="list-style-type: none"> <li>Interviewees' perception.</li> <li>Funding agencies' statements.</li> <li>Media.</li> </ul>
Novel quality control	<ul style="list-style-type: none"> <li>Which novel modes of quality control have emerged?</li> <li>Do these replace or complement older forms?</li> <li>Are researchers aware of and comply with different modes?</li> <li>What actors represent these new modes?</li> </ul>	<ul style="list-style-type: none"> <li>Interviewees' perception.</li> <li>Professional standards.</li> <li>Funding rationales and streams from various agencies.</li> </ul>

policy domains (Zapp and Powell 2016). By the late 1990s, Anglophone countries already experienced their 'evidence turn' (Banks 2009; Howlett 2009). German policy-makers' interest in educational data and its applicability is evident in all relevant statements, explaining most of the recent funding enthusiasm in ER. The BMBF sees national and international assessments, the search for causal mechanisms driving educational system development, educational statistics, and a 'system of internal and external evaluations' as directly linked to evidence-based practices enabling 'long-term preventive policy-making' (BMBF 2008: 6; also WR 2001: 69). Under the German EU Presidency, the BMBF invited 300 members from the European Union and Council of Europe member countries to the first Europe-wide conference on 'Knowledge for Action. Research Strategies for an Evidence-based Education Policy' (2007). The State Ministries of Education organized in the Standing Conference of Ministers of Culture, or *Kultusministerkonferenz* (KMK 2006: 5), too, is convinced that 'outcome-orientation, accountability and system monitoring mark a paradigm change' in German education policy. Research-based knowledge becomes the primary basis for making political decisions: 'As with other policy domains: political action in education can only be genuinely responsible if we face and take into account scientific knowledge and findings' (Federal Minister of Education, Annette Schavan 2009: 3).

Importantly, the new emphasis on evidence for policy-making accompanied the introduction of educational reporting mechanisms.

BMBF and KMK agreed on comprehensive educational monitoring (KMK 2006) with the first National Education Report published in 2006 and appearing biannually since, with each report focusing specific topics for presentation in-depth, such as education and migration (Autorengruppe Bildungsberichterstattung 2006, 2016) or inclusive education (Autorengruppe Bildungsberichterstattung 2014). In parallel, state educational reports (produced by more than ten *Länder* by 2015) and numerous local or district educational reports increase available knowledge on educational developments (see Busemeyer and Vossiek 2015).

Such a turn toward more systematic educational reporting requires a regulatory and empirical basis of indicators; these proliferate in an era of big data. National educational standards now used as benchmarks in *mandatory* intra-national comparative studies were introduced in 2003. States also agreed to make participation in certain ILSAs *mandatory* (TIMSS, PIRLS/IGLU, PISA). The National Educational Panel Study (NEPS), launched in 2009, is a highly significant data generator across the life course and at all levels of learning that explicitly aims to inform educational reporting and policy advice (e.g. Leuze (2008)).

Policy-makers' interest in knowledge usability is reflected in massive investments in the infrastructure necessary to conduct the envisioned EER. Hefty research programs such as the *Framework Program for the Promotion of Empirical Educational Research* represent an unprecedented intervention in research on education in

**Table 2.** Educational research funding priorities by source, 1998–2007

	Type of research						Total (%)
	Applied	Descriptive	Empirical	Basic/methodology	Practical/informative/ documentation	Miscellaneous	
DFG	9.1	7.9	49.4	5.4	14.7	13.6	559 (22)
BMBF	57.2	4.7	26.5	2.3	4.2	5.1	215 (8.5)
EU	66.7	4.5	18.7	0.5	1.0	8.6	198 (7.8)
Other	42.7	4.8	35.7	2.6	4.6	9.6	1,535 (61)

Source: Weishaupt and Rittberger (2012); partly own calculations; percentage of funding by funding source.

financial, structural, and substantive terms. Between 2007 and 2019, the Framework Program funds roughly 342 research projects in eleven defined priority areas, including a data center, a longitudinal panel study (NEPS), and the training of hundreds of young researchers in EER. The total sum—over €163m—is exceptional, especially considering that educational departments within German higher education institutions receive between €45m (2007) and €80m (2013) annually.<sup>1</sup> Other important initiatives include the higher education funding line (2008–18; €90.2m; 154 projects),<sup>2</sup> state-by-state comparisons of school achievement, and international assessments, such as the OECD's triennial PISA, for €15m per round.

Researchers that can be seen as being among the main winners of these investments stress that they see science as morally obligated to return something to society, to repay its debts in research understood to be *relevant* (A VII). Such an understanding clashes with beliefs held by representatives from traditional educational scholarship. In fact, influential traditional DGfE members accuse the new generation of 'methodological economism' (Radtke 2015), the latter their pedagogical counterparts of 'other-wordly aestheticism' (A III), even as they advocate and conduct research with a 'service character' (A II, III, IV, VII, VIII). Both sides argue for the importance of their approach, with few boundary-spanners able to synthesize and argue for combinations.

Beyond this methodological and paradigmatic turf battle,<sup>3</sup> numbers confirm the trend toward more 'empirical' material and 'finalization' of German educational science, aided by agenda-setting and priorities in funding by certain actors. First, professors with a clear EER profile are the biggest winners in recruitments over the past decade. Their number tripled in the period 2006–12 alone (see below). Second, ever-more important external funding from the federal BMBF and the EU lean toward projects with an 'applied' component (Table 2). Two thirds of European funding and well over half of federal funding is channeled toward research with an explicit 'applied' label, while the DFG (traditionally the primary funding agency) directs less than 10 per cent of its funds to that kind of research, prioritizing fundamental research.

## 6.2 Multi-/transdisciplinarity

In assessing the degree of trans-/multidisciplinarity within this kind of ER burgeoning in Germany, we first focus on the self-portrayal of its professional representation and the perceptions of key actors involved therein. Two contrasting associations represent the ER community in Germany, with the sixty-year-old DGfE now joined by German Empirical Educational Research Association (*Gesellschaft für Empirische Bildungsforschung*, GEBF). Founded in 2012 by influential educational researchers to dissociate themselves from humanities-oriented educational science, the GEBF embraces multi- and transdisciplinarity.

Interviewees involved in GEBF compare it to problem-oriented scholarly communities in climate research, maritime research, public health, or the life sciences (A III, VII, VIII). They stress the GEBF's transdisciplinary nature that comes with such issue-focused work, enabled by usage of similar methods. Deliberately, it seeks to attract researchers from diverse education-related fields, such as psychology, sociology, economics, statistics, and computer science—so long as they share a common *empirical* claim. They confirm that their new research infrastructure revolves around an 'object' (A V) or an 'issue' (A VIII) more so than a discipline. Importantly, the issues this new scientific infrastructure seeks to address emanate less from the scientific system itself, rather than the practice, the field, and the society. GEBF's annual conferences usually display transdisciplinarity in their calls and the invited contributions reflect this claim. Proceedings of GEBF's second annual conference in 2014 begin with the following words: 'The disciplinary and institutional fragmentations of science are not natural, but historical products. The increasing complexity of current research questions requires responses that assume inter- and transdisciplinarity, from basic research to its application' (GEBF Conference Proceedings 2014: 7).

The EER research community, due to its complex organizational network structure, is proud to proclaim that it has the

huge potential to overcome the kind of education research that was confined to disciplinary boundaries until now and to do what is nowadays crucial in many domains: acting problem-focused and research field-based, which always means to span and to transgress disciplines (A III).

We are witnessing times in which people think big and in interdisciplinary terms. If you look at where policymakers direct their funding, you can see, the bigger, the better, the more interdisciplinary, the better. The hotter the topic, the better. Education is a hot topic and a topic that cannot be researched by one single discipline individually. [...] This is the research model of the near future since such enterprises have policymakers' ears (A VIII).

A second kind of evidence supports the Mode 2 claim of increasing multidisciplinarity, the systematic collaboration of more than two disciplines. Here, we draw on our analysis of the (number of) disciplines involved in ER projects ( $N = 150$ ). While the data does not allow us to trace an assumed increase in multidisciplinarity across time we can compare the degree of multidisciplinarity in DFG-funded basic ER as opposed to BMBF-funded 'applied' ER in its two major funding lines, the Framework Program and the Higher Education Research Program (see above). For both DFG and BMBF projects, almost half of all projects include more than one discipline. Interestingly, BMBF projects have an average of 2.6 disciplines participating, while DFG projects unite only 1.5 disciplines, perhaps due to the disciplinary constitution of DFG peer review. Education science (or pedagogy) participates in around a third of these DFG projects as lead discipline, whereas in BMBF projects only around a

quarter. In all, 22 per cent are active in the areas of psychology and psychiatry, followed by subject didactics (8.5 per cent) and language and literatures (6.7 per cent).<sup>4</sup> Remarkably, if the funding line in higher education is analyzed separately, pedagogy is largely outstripped by any other discipline, emphasizing the shift in naming from 'education science' to the inclusive 'educational sciences'.

### 6.3 Organizational diversity

Our German ER organizational network analysis uncovers a sprawling landscape cutting across categories of governance levels (international/national/subnational), organizational types (public/private), missions (scientific/commercial), functions (primarily knowledge-producing/knowledge-using), and disciplines (e.g. education, psychology, sociology, history, neuroscience, economics, computer science, etc.). Certainly, ties are tighter between national organizations and those organized within the same umbrella association (e.g. the German Leibniz consortium Bildungspotenziale, LERN). Yet international organizations also routinely interact with these national organizations, including the Australian Council for Educational Research (ACER), the non-governmental Cochrane Center, the commercial Educational Testing Service (ETS), and the Data Processing Center (DPC) at the International Association for the Evaluation of Educational Achievement (IEA). These are international research or research-related institutes specialized in educational test design, assessment, analysis, and consulting. They are highly intertwined with national research institutes. Large-scale assessments, such as OECD's PISA, for instance, require expertise from IEA and ACER. Leibniz institutes (e.g. IPN and DIPF) in charge of PISA contract DPC to carry out data collection, which, in turn, borrows sampling strategies from North America: 'Certain research institutes have specialized in a particular field, they do what they can do best. In other cases, as with large-scale sampling, there are only a few institutes worldwide that can do that' (A IX). Another example is Bamberg's Leibniz Institute for Educational Trajectories (LIfBi), which, in designing the prestigious NEPS drew on experience from the US ETS (recruiting staff) (Schavan 2009).

If many of the organizations have been active for decades, it is not without a hint of irony that some of these institutes were at the brink of being shut down before EER took hold, literally giving them a new lease on life (A II, VII, VIII, IV). Now displaying considerable expansions in terms of operational portfolios, research units, budgets and staff, Leibniz institutes like IPN and DIPF have doubled their staff in just the last seven years to more than 170 and 300 employees, respectively. The IEA's DPC, founded in 1995 counting 5 people, is now staffed with 130 employees.

Universities are somewhat conspicuous by their absence in leading this field. To be sure, overall, universities account for the biggest share in producing and using ER and they carry out 80 per cent of all ER projects compared to 9 per cent at extra-university institutes. Yet, these institutes are the big winners of the last fifteen years, as they conduct policy-relevant, applied EER and far more contract research than do universities (38 per cent versus 21 per cent). This is reflected in the commitment of their most important funder, the federal BMBF, providing a third of their funding (universities only 14 per cent) (Faulstich-Wieland 2012; Botte et al. 2012). Moreover, institutes such as those organized within the Leibniz or Max Planck Associations are far more research-intensive, as they dedicate 13 per cent of their total R&D expenditures on education, while universities with their more diverse missions spend only between 2.5 per cent and 3 per cent on average (Botte et al. 2012).

Only recently have members of the field begun to more strongly link their activities with the higher education system. Some organizations have affiliated with universities, sometimes due to prominent individuals continuing their ER career in universities. Strong legitimacy, generous long-term funding, and a fertile research agenda assured by abundant data have impacted higher education. If job postings in EER more than doubled from 2003–6 to 2007–10 (from 34 to 73), the actual number of university professors in EER almost tripled from 27 (3 per cent of all education professors) in 2006 to 75 (8.5 per cent) in 2012 (Aljets 2015), ensuring future growth in EER.

### 6.4 Reflexivity and accountability

To understand the new sense of reflexivity embraced by researchers, the GEBF is crucial. Its founders' stated goals not only include the collaboration of diverse disciplines and the publication and dissemination of research findings, but also the *transfer* of the state of the art and the development of EER to educational practitioners *and* the public (GEBF 2012). Note that authors aim to communicate the *situation of the field as a whole* to audiences far beyond the scientific community. Interviewees involved in GEBF describe it as an essential part of their day-to-day work to communicate with representatives from 'federal, state and district level educational policymaking and administration' (A VII), to engage in 'knowledge transfer' with different stakeholders, especially political decision-makers that need systematically reviewed syntheses (A III, V), and to convince school districts to participate in studies (A VIII). In general, prominent individual researchers are consulted worldwide to assist governments in designing research programs (A X).

The aforementioned important BMBF-funded Framework Program explicitly tasks affiliated researchers to publish their research findings in 'widely understandable forms', to present their 'scientific questions and results within the framework of a societal discussion', to prepare results for scientists and a wider audience, to reach out to practitioners, and to establish ER as a key issue in the public discourse (BMBF 2008: 20).

Another form of contextualization occurs at the level of policymaking. Between 2009 and 2013, Federal Ministries in Germany paid €1 billion for external consultants, with the BMBF alone accounting for almost half of that sum (€465m). In the long list of its expert clients, most of the institutes mentioned above benefited (Antwort der Bundesregierung auf die Kleine Anfrage Drucksache 17/14370—Bundestag, 17. Wahlperiode, August 2013).

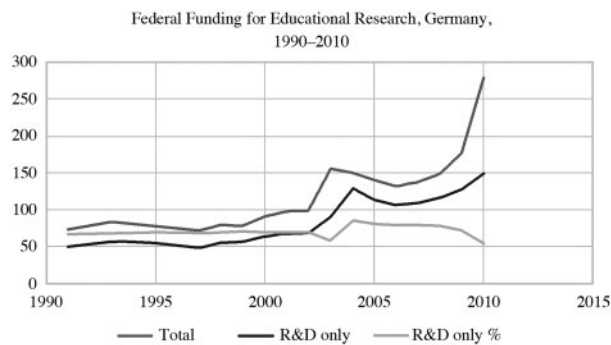
Yet, contextualization also works in the opposite direction, that is, from public stakeholders to educational expertise. One telling demonstration of how this new kind of accountability influences the emerging ER is how roles and approaches are discussed in the mass media. Ties Rabe, former president of the *Kultusministerkonferenz* and Olaf Köller, current GEBF president, debated the value of EER and its potential for policy advice in general and the role of large-scale assessment in particular in Germany's influential newsweekly *Die Zeit* (January 2013; Nr. 4).

Furthermore, the results of the first PISA round in 2000, unexpectedly showing below average results for German 15-year olds, precipitated a crisis of legitimacy (Waldow 2009). Suddenly, teachers, schools, families, researchers, and policy-makers became involved in heated debates on German education and ER, fueled by extraordinary interest from the mass media. PISA repeatedly makes headlines in all major German newspapers, renewing the deep soul-searching last witnessed in the 'educational catastrophe' (*Bildungskatastrophe*) of the 1960s (Picht 1964), with the German

**Table 3.** Sources of Educational Research Funding in Germany, 1998–2014 (million €)

	1998–2003 (N = 853)	2003–05 (N = 500)	1998–2007 (N = 1,701)	2011–14 (N = 4,160)
DFG	14.4	17	22.2	40.7
BMBF	8	9.8	8.5	16.0
EU	4.6	5	7.8	4.8
Other (philanthropies, state and federal agencies, DAAD)	73	68.2	61.5	38.5
Total	100	100	100	100

Sources: Datenreport 2004–16.



**Figure 1.** Federal Funding for Educational Research, Germany, 1990–2010 (million €)

Source: Muders and Weishaupt (2012).

Education Ministries declaring an ‘educational emergency’ (*Bildungsnotstand*) (Powell and Solga 2006). Every PISA round is now highly mediatized in the printed press, television, radio, and online—to an extent found in hardly any other participating country (Waldow 2009; Gläser et al. 2014; Aljets 2015).

### 6.5 Novel quality control

Quality control under Mode 2 conditions is said to shift from science-based peer review to a wider audience. Thus, it may be somewhat surprising that the rise of EER brought with it a clear mandate for high-quality, peer-reviewed scientific research publications. These challenged the ‘old generation of educational researchers content with writing aesthetic essays with a real quality problem’ that dominated humanities-based education scholarship in Germany until the 1990s (A III). The ‘new’ educational research generation is proud to display its excellence by high methodological standards (randomized control trials, quasi-experiments, systematic reviews) and international scope through English-language (article) publications.

However, a second change is then reflected in German funding structures. Federal and state policy-makers have started to boost EER with unusually large and direct financial measures. Traditionally, the German Research Foundation (DFG), which, although completely dependent on generous federal funding, acts as a highly-independent allocator of resources in all scientific disciplines, including both university and nonuniversity research, based on extensive reviews of scientific quality. This has changed with the start of the above-mentioned Framework Program and other federal initiatives. Several observers (A II, VII, VIII) noted that this marks a fundamental change in funding policy since state-commissioned research of such scale is unprecedented in Germany. More importantly, the funding logic of the BMBF differs from the DFG’s in that

it puts far more emphasis on *economic and social impact than scientific quality* alone, even if the rigorous DFG peer reviews are emulated. This new paradigm advances a ‘utilitarian rationale for conducting ER.

What has been standard procedure in many other countries and in EU-funded Framework Programs, has only now made its way into German national research policy (Table 3). Educational policy-makers themselves feel that such rapid and massive influx of investments is necessary if ‘you want to strategically develop a research field that combines excellence with a service mandate for policy-making’ (A I, VI). As seen in Table 1 in Section 6.1 above, the type of research funded (e.g. fundamental vs. applied) varies according to funding agency. Most direct federal and EU funding is channeled toward research that is explicitly labelled ‘applied’, while DFG resources are not. Such a premium on ER with an applied focus might increase given the growing attention to ER at the federal level indicated by a strong increase in funding over the past ten years (Fig. 1). Furthermore, the EU, too, has found its place in the funding portfolio and displays a clear preference for applied research.

## 7. Discussion and conclusion

Composed of scholars from various disciplines, beyond education, including sociology, psychology, and economics, members of the field of EER in Germany deliberately seek proximity to policy-makers and assume a decidedly international profile. The field has integrated a wide array of organizational types, many outside of the university. Strongly anchored in increasingly routinized international large-scale assessments, the field is thus embedded in wider global networks of education researchers. These researchers are active in research institutes and think tanks and well-connected to policy-makers. The field reflects strong applicability to its context, is explicitly trans- and multidisciplinary, exhibits organizational diversity, reflexivity and accountability, and has started to integrate novel quality criteria. Thus, the empirical case of the new generation of ER in Germany displays, to certain degrees, all of the transformative elements predicted by the NPK theory. Based on the data presented in the case study of German ER, the tenets of Mode 2 facilitate analysis of the changing conditions for ER. While the transformation analyzed is neither pure nor complete, the operationalization of Mode 2 concepts supported comprehensive discussion of complex change processes in a diverse field in which different scholarly strands continue to coexist.

We discuss aspects that address key criticisms levelled against the framework. Criticism of the alleged ‘novelty’ of Mode 2 (Ertzkowitz and Leydesdorff 2000; Rip 2000) does not apply to the German situation. Here, indeed, the dominant approach to ER until the 1990s was Mode 1. More specifically, Godin (1998) defends the claim that the social sciences have always been, to some extent,



operating in Mode 2. The national particularities of education scholarship in Germany prior to the observation period have, however, included its distance to applicability, its mostly academic organization in disciplinary isolation, and its very low quality control standards (albeit not necessarily low quality) as compared to the contemporary situation. In stark contrast to that traditional model, the new EER infrastructure is organizationally heterogeneous. Its members are committed to thorough internationally-recognized scientific quality control and seek legitimacy by conducting ‘socially relevant’ ER. Tellingly, not only does this field offer its knowledge, but it also has a clear understanding of the kinds of knowledge it offers. Interviewees distinguished between three different knowledge types: knowledge that identifies problems; knowledge that systematizes the state of the art (systematic reviews, syntheses); and knowledge for action and implementation. Most of the involved researchers believe that all of these types are successfully being provided to policy-makers and practitioners (A II, III, IV, VII, VIII).

Moreover, contrary to Kropp and Blok’s (2011) account of Danish sociology, change was discontinuous, even disruptive, following revelatory international large-scale assessments and a heightened interest in ‘high-quality’ ER among policy-makers and the wider public. Large sections of preexisting research infrastructure have been converted to a Mode 2 mission and additional research infrastructure established is clearly compatible with Mode 2, namely as a scientific service sector. While prior structures supporting empirical ER were present, they were less central, even scattered. Expanding and assembling these disciplines and actors is among the main rationales of EER.

Further, while Shinn (2002) rightly asserts that ‘university, business and government all function in a national setting’, he discounts the organizational heterogeneity in research operations, with new actors, changing portfolios and scopes, emerging ad-hoc collaborations and quasi-permanent networks that span borders of type, mandate, and geography. In this sense, it is a global, porous, and permeable research field in the making. Thus, it is telling that some of the founders of the GEBF, when speaking about their organization, actually use Gibbons and Nowotny’s (2001: 69) language of a ‘forum’ or ‘platform’ (A VII, VIII) to describe the rationale of their multi- or transdisciplinary association.

The international outlook is a particularly strong theme in contemporary German ER. On numerous occasions, policy-makers and advisory bodies stressed the importance of ‘catching up’ and ‘keeping up’ with ‘international knowledge standards’ (WR 2001: 72) situating Germany in the context of the ‘European educational space’ (Schavan 2009: 2). Internationalizing ER was the only way to overcome the ‘provincialism’ (Actor III) of Germany’s traditional educational scholarship (see also Tröhler (2014)).

One reason that such rapid change and innovation mainly occurs outside of academia has been given by organizational leaders themselves: Not only are visionary leaders taking their organizations to necessary new scientific frontiers, but also guaranteeing the room for maneuver or organizational reactivity to do so, something less often found within teaching-oriented academic departments (A V, VI). This judgment reflects the Mode 2 argument that the ‘emergence of loose organizational structures, flat hierarchies, and open-ended chains of command’ represent most fertile ground for future innovations (Gibbons and Nowotny 2001: 69; also Gibbons et al. 1994: 7).

Importantly, in identifying the causes for such institutional change, it would be shortsighted to restrict the focus to endogenous, national and functionalist explanations, as do Kropp and Blok (2011)

and Gibbons and Nowotny (2001). Our case illustrates how *international* stimuli, in the guise of large-scale assessments, evidence-based policy-making experiments, and a globe-spanning quantitative research paradigm in general trigger cognitive and organizational changes in a national setting (Zapp and Powell 2016). Such large-scale shifts are supported by a highly utilitarian science for development policy model institutionalized on highly collective levels and beyond concrete evidence of its viability (Drori et al. 2003).

The case of Germany not only sheds light on the construction and organization of a Mode 2 research infrastructure, but also points to several threats, something much less discussed in the original contribution or subsequent reactions to it.

A first risk concerns the *organizational* or *management structure* of Mode 2 ER. Many of our interviewed experts look at this sprawling international and interorganizational program with mixed feelings. Some marvel at the complex, but smooth operation of large-scale assessments and national follow-ups (A VII). Others point out the monopolies created in data collection and analysis (A V). Some feel that decision-making atomizes within these larger programs in which a ‘holy alliance between science, politics, IOs and industry’ (A VIII) takes shape, resulting in scientists in particular locales having little say. The competence measurement boom, for example, risks inviting purely commercial actors, such as large IT firms, to enter the field with their research portfolios, edging out scientific priorities and paradigms (A VIII).

A second threat touches on the *relationship between research and non-scientific stakeholders*. It is not surprising that in the highly contextualized domain of the common good of education, parents’ associations, teacher unions, school leaders as well as district, state and federal administrators alike act as stakeholders that have something to say about how education could or should be researched. Also, the proximity to policy-makers is particularly ambivalent. Polemic reactions from the scholarly community await those researchers that move too close to policy-making debates (A VII). Some interviewees call it an ‘unfortunate incident’ (A V, VII) that the notion of evidence-based policy-making has become increasingly influential, as they fear that policy-makers’ expectations of ‘hands-on research findings’ (A V) have been unduly fueled. More realistic prospects for innovation could well lower research policy generosity in the future and even call the EPM experiment into question.

A third, and related, threat emanates from new forms of *quality control* heralded by the Mode 2 authors, yet critically judged by many interviewees in our study. Large research programs, such as the Framework Program or the Higher Education Funding Priority, shift the mode of governance from structural aspects to a more epistemic approach (i.e. defining priority areas). Some observers fear threats to research quality, viewing these investments as ‘easy and cheap money’ for what is temporarily *en vogue*, a passing political whim (A II, AVIII) and calling for more careful spending in the future with a stronger emphasis on quality criteria in considering research proposals (A V, VIII).

Considering that both BMBF and the EU established crucial roles in research funding for ER over the past two decades and both emphasize the service function of research, there is tangible reason to believe that this will have a lasting impact on the rationales of research projects and the cognitive development of the field as a whole (Edler and Kuhlmann 2008; Kastrinos 2010).

Finally, our data suggest an additional aspect that has not been given particular attention in the Mode 2 or related literature. This concerns the *sustainability* of Mode 2. A distinct feature of EER in Germany has been and continues to be its strong program-based

foundation. The federal programs for EER and HE as well as numerous (partly mandatory) international assessment programs (twelve since 2000) entail a massive increase in project work, with resulting data piles providing (not only dissertation) fodder. Key organizations rely heavily on program- and project-based funding, with their scientific and administrative staff members working largely within this framework. At DIPF and IPN, for example, two of the largest ER organizations in Germany, more than half of the scientific staff are employed on untenured project-basis (161 of 300 at DIPF; 93 of 170 at IPN).

Yet, the ongoing dominance of expensive, large-scale, program-based, and project-laden funding may not be sustainable. If governmental action has always been important in creating and shaping certain research fields due to the necessary investments (Cozzens and Woodhouse 1995), in Germany this is a relatively new phenomenon. It demonstrates the paramount importance currently conferred on education. Simultaneously, actors complain about the vicissitudes of the research agenda, with months of relentless work followed by times of labor slack (A IX). Some forecast an inescapable shake-out in store for the research field, fearing that too many empirical researchers, trained and recruited in large international and national programs, will face gloomy job prospects if resource levels decline, particularly in a state-dominated higher education system in which tenured and professorial positions are scant (A V, VI). Still others see research priorities, if defined top-down, as reflecting political agendas more than scientific rigor (A IX). Thus, the contemporary era is characterized by an extraordinary, albeit ambivalent growth of an entire research infrastructure united by a new paradigm that reflects the dimensions unified in the NPK framework.

Applied here, the Mode 2 concept as an analytical framework enabled us to trace changes in ER that have occurred in Germany since 2000. German ER has shifted considerably away from its heritage as introverted, university-based, and policy-decoupled disciplinary scholarship, rooted mainly in the humanities. Today, an increasing portion of educational scholarship considers itself ‘empirical’ and relies on an open, organizationally-heterogeneous, and service-oriented research infrastructure built primarily with government funds to process vast quantities of largely quantitative data.

Utilized as a heuristic device, Mode 2 uncovered fundamental changes in HE and research systems, key drivers thereof, and diverse institutional and organizational implications. However, to do so, we needed to further refine the analytical apparatus provided by its proponents. In general, we encourage future studies to further empirically test the hypotheses of the NPK. Indeed, future studies may apply the concepts of Mode 2 knowledge production to other scientific fields, allowing researchers to evaluate long-term trends in research systems. Conversely, if it is not to remain a largely theoretical artefact, the conceptualization of Mode 2 science may well benefit from being enriched by empirical insights deriving from analyses framed more carefully in specific disciplinary, cultural, and historical contexts.

## Notes

1. The year 2013 is the latest data available; taken from *Datenreport Erziehungswissenschaften* (2016). We thank Bernhard Schmidt-Hertha for providing most recent funding data ahead of publication.
2. Information provided by the project-coordinating DLR in an email from 28 July 2015.
3. In a peculiar analogy to the so-called ‘methodological

dispute’ (*Methodenstreit*) in the Germanophone social sciences in the 1960s between positivists and critical theorists, two paradigms seem to clash again in education, bringing the debate into policy-making agencies and spilling over into the mass media (see below).

4. See Schmidt-Hertha (2012: 170) for a similar count in BMBF Framework Program projects.
5. See Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Dr Barbara Höll, Harald Koch, Richard Pitterle, Dr Axel Troost und der Fraktion DIE LINKE—Drucksache 17/14370—Ausmaß der Vergabe von Aufträgen der Bundesregierung an externe Dritte. Deutscher Bundestag, 17. Wahlperiode. 28 August 2013.

## References

- Albert, M. (2003) ‘Universities and the Market Economy: The Differential Impact on Knowledge Production in Sociology and Economics’, *Higher Education*, 45/2: 147–82.
- Aljets, E. (2015) *Der Aufstieg der Empirischen Bildungsforschung. Ein Beitrag zur institutionalistischen Wissenschaftssoziologie*. Wiesbaden: Springer VS.
- Autorengruppe Bildungsberichterstattung (2006–2014) *Bildung in Deutschland*. Bielefeld: Bertelsmann.
- Baker, D. (2014) *The Schooled Society*. Stanford: Stanford University Press.
- Banks, G. (2009) Evidence-Based Policy Making: What it is? How do we get it? ANU, Canberra, Australia. Public Lecture Series, presented by ANZSOG. <[http://www.pc.gov.au/\\_data/assets/pdf\\_file/0003/85836/cs20090204.pdf](http://www.pc.gov.au/_data/assets/pdf_file/0003/85836/cs20090204.pdf)> accessed 15 Dec 2015.
- Beck, U., Giddens, A., and Lash, S. (1994) *Reflexive Modernization*. Stanford: Stanford University Press.
- Benavot, A., and Meyer, H. D. (2013) *PISA, Power, and Policy: The Emergence of Global Educational Governance*. Oxford: Symposium Books.
- Biesta, G. (2015) ‘On the Two Cultures of Educational Research, and How We Might Move Ahead: Reconsidering the Ontology, Axiology and Praxeology of Education’, *European Educational Research Journal*, 14/1: 11–22.
- Botte, A., Dees, W., Huth, R., Martini, R. et al. (2012) *Bildungsforschung in Deutschland. Eine Situationsanalyse*. Frankfurt am Main: Bundesministerium für Bildung und Forschung.
- Bogner, A., Littig, B., and Menz, W. (2009) *Interviewing Experts*. Houndmills: Palgrave Macmillan.
- Bundesministerium für Bildung und Forschung (2008) *Framework Program for the Promotion of Empirical Educational Research [Rahmenprogramm zur Förderung der empirischen Bildungsforschung]*. Bonn/Berlin: BMBF.
- Busemeyer, M. R., and Vossiek, J. (2015) *Reforming Education Governance Through Local Capacity-building*. OECD Education Working Papers No. 113. Paris: OECD.
- Cozzens, S. E., Woodhouse, E. J. (1995) ‘Science, Government, and the Politics of Knowledge’. In: Jasanoff, S., Markle, G.E., Petersen, J.C. et al. (eds) *Handbook of Science and Technology Studies*, pp. 533–53. London: SAGE.
- Crompton, H. (2007) ‘Mode 2 Knowledge Production: Evidence from Orphan Drug Networks’, *Science and Public Policy*, 34/3: 199–211.
- Dederig, K. (2009) ‘Evidence-based Education Policy: Lip Service or Common Practice? Empirical Findings from Germany’, *European Educational Research Journal*, 8/4: 484–96.
- Deutsche Forschungsgemeinschaft (2002) ‘Stellungnahme zur strukturellen Stärkung der empirischen Bildungsforschung’, *Zeitschrift Für Pädagogik*, 48/5: 786–98.
- Drori, G. S. et al. (2003) *Science in the Modern World Polity*. Stanford: Stanford University Press.
- Edler, J., and Kuhlmann, S. (2008) ‘Coordination within Fragmentation: Governance in Knowledge Policy in the German Federal System’, *Science and Public Policy*, 35/4: 265–76.

- Elzinga, A. (1985) 'Research, Bureaucracy and the Drift of Epistemic Criteria'. In: Wittrock, B., Elzinga, A. (eds) *The University Research System*, pp. 191–220. Stockholm: Almqvist & Wiksell International.
- (1997) 'The Science-Society Contract in Historical Transformation', *Social Science Information*, 36/3: 411–45.
- Etzkowitz, H., and Leydesdorff, L. (1998) 'The Endless Transition: A "triple helix" of University–Industry–Government Relations', *Minerva*, 36: 203–8.
- , and — (2000) 'The Dynamics of Innovation: From National Systems and "Mode 2" to a Triple Helix of University–Industry–Government Relations', *Research Policy*, 29/2: 109–23.
- Faulstich-Wieland, H., Horn, K. P., Thole, W. et al. (eds) (2012) *Datenreport Erziehungswissenschaft*. Opladen: Verlag Barbara Budrich.
- GEBF (2012) Satzung der Gesellschaft für empirische Bildungsforschung. <<http://www.gebf-ev.de/%C3%BCber-die-gebf/satzung/>> accessed 6 Jan 2016.
- (2014) Conference Proceedings. <<http://www.gebf2014.de/>> accessed 20 Jan 2016.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S. et al. (1994) *The New Production of Knowledge*. London: SAGE.
- , and Nowotny, H. (2001) 'The Potential of Transdisciplinarity'. In: Klein, J. T., Grossenbacher-Mansuy, W., Häberli, R., Bill, A. et al. (eds) *Transdisciplinarity: Joint Problem Solving among Science, Technology, and Society*, pp. 67–80. Basel: Birkhäuser.
- and — (2003) 'Introduction: "Mode 2" Revisited: The New Production of Knowledge', *Minerva*, 41/3: 179–94.
- Gläser, J., Aljerts, E., Gorga, A., Hedmo, T., Håkansson, E. and Laudel, G. (2014) 'Path Dependence and Policy Steering in the Social Sciences', *Research in the Sociology of Organizations*, 42: 267–95.
- Godin, B. (1998) 'Writing Performative History: The New New Atlantis?', *Social Studies of Science*, 28/3: 465–83.
- Hamann, J. (2015) 'Die Geisteswissenschaften und ihr Bildungsdiskurs', *Zeitschrift Für Soziologie*, 44/3: 180–96.
- Hessels, L. K., and van Lente, H. (2008) 'Re-thinking New Knowledge Production: A Literature Review and a Research Agenda', *Research Policy*, 37: 740–60.
- Heyneman, S. P., and Lykins, C. (2008) 'The Evolution of Comparative and International Education Statistics'. In: Ladd, H. F. and Fiske, E. B. (eds), *Handbook of Research in Education Finance and Policy*, pp. 105–127. New York: Routledge.
- Hicks, D. M., and Katz, J. S. (1996) 'Where is Science Going?', *Science Technology and Human Values*, 21/4: 379–406.
- Horlacher, R. (2016) *The Educated Subject and the German Concept of Bildung*. Abingdon: Routledge.
- Howlett, M. (2009) 'Policy Analytical Capacity and Evidence-based Policy-Making: Lessons from Canada', *Canadian Public Administration/ Administration Publique Du Canada*, 52/2: 153–75.
- Ingenkamp, K. (1992) 'Die Anfänge der Arbeitsgruppe für empirische pädagogische Forschung, AEPF', *Empirische Pädagogik*, 6/1: 109–17.
- Kastrinos, N. (2010) 'Policies for Co-ordination in the European Research Area', *Science and Public Policy*, 37/4: 297–310.
- Keiner, E. (1999) *Erziehungswissenschaft 1947-1990*. Weinheim: Deutscher Studien Verlag.
- Kerstan, T., and Spiewak, M. (2013) Ranglisten sind gefährlich. Die Zeit, Nr. 4, January 2013.
- Kleine Anfrage der Abgeordneten, Barbara, H., Harald, K., and Richard, P., Axel Troost und der Fraktion DIE LINKE—Drucksache 17/14370—Ausmaß der Vergabe von Aufträgen der Bundesregierung an externe Dritte. Deutscher Bundestag, 17. Wahlperiode. 28 Aug 2013.
- Krohn, W., van den Daele, W., and Weingart, P. (eds) (1979) *Geplante Forschung*. Frankfurt/Main: Suhrkamp.
- Kropp, K., and Blok, A. (2011) 'Mode-2 Social Science Knowledge Production? The Case of Danish Sociology between Institutional Crisis and New Welfare Stabilizations', *Science and Public Policy*, 38/3: 213–24.
- Kultusministerkonferenz (1997) Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland. Konstanzer Beschluss. <[http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/1997/1997\\_10\\_24-Konstanzer-Beschluss.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf)> accessed 11 May 2016.
- (2006) Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. <[http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2006/2006\\_08\\_01-Gesamtstrategie-Bildungsmonitoring.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2006/2006_08_01-Gesamtstrategie-Bildungsmonitoring.pdf)> accessed 11 May 2016.
- Leuze, K. (2008) *Bildungswege besser verstehen: das Nationale Bildungspanel*. WZBrief Bildung 02. Berlin: WZB.
- Leydesdorff, L., and Meyer, M. (2006) 'Triple Helix Indicators of Knowledge-based Innovation Systems: Introduction to the Special Issue', *Research Policy*, 35/10: 1441–9.
- Muders, S., and Weishaupt, H. (2012) 'Finanzierung der Bildungsforschung'. In: Weishaupt, H. and Rittberger, M. (eds) *Bildungsforschung in Deutschland: eine Situationsanalyse*. Frankfurt am Main: Bundesministerium für Bildung und Forschung.
- Nowotny, H., Scott, P. and Gibbons, M. (2001) *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge: Polity Press.
- Pestre, D. (2003) 'Regimes of Knowledge Production in Society', *Minerva*, 41: 245–61.
- Picht, G. (1964) *Die deutsche Bildungskatastrophe*. Olten: Walter.
- Powell, J. J. W., and Solga, H. (2006) 'Gebildet/Ungebildet'. In: Lessenich, S., Nullmeier, F. (eds.) *Deutschland. Eine gesplante Gesellschaft*, pp. 175–90. Frankfurt am Main: Campus.
- Radtke, F. O. (2015) 'Methodologischer Ökonomismus', *Zeitschrift Für Erziehungswissenschaft*, 26/50: 7–16.
- Rip, A. (2000) 'Fashions, Lock-Ins and the Heterogeneity of Knowledge Production'. In: Jacob, M. and Hellström, T. (eds), *The Future of Knowledge Production in the Academy*, pp. 28–39. Buckingham: SRHE, Open University Press.
- Roth, H. (1963) 'Die realistische Wendung in der pädagogischen Forschung', *Die Deutsche Schule*, 55: 109, 119.
- Schavan, A. (2009) Zur Bedeutung von Bildungsforschung für eine wissenschaftsbasierte Politik und Praxis. Speech at Inaugural Ceremony of NEPS.
- Schmidt-Hertha, B. (2012) 'Forschung und Publikationskulturen'. In: Thole, W., Faulstich-Wieland, H., Horn, H.-K., Weishaupt, H. and Züchner, I. (eds), *Datenreport Erziehungswissenschaft*, pp. 159–180. Opladen: Barbara Budrich.
- Shinn, T. (1999) 'Change or Mutation? Reflections on the Foundations of Contemporary Science', *Social Science Information*, 38/1: 149–76.
- (2002) 'The Triple Helix and New Production of Knowledge', *Social Studies of Science*, 32/4: 599.
- Slaughter, S. and Rhoades, R. (2009) *Academic Capitalism and the New Economy*. Baltimore: The Johns Hopkins University Press.
- Slaughter, S., and Leslie, L. L. (1997) *Academic Capitalism*. Baltimore: The Johns Hopkins University Press.
- Terhard, E. (2016) 'Interdisciplinary research on education and its disciplines: Processes of change and lines of conflict in unstable academic expert cultures: Germany as an example', *European Educational Research Journal*, 1–16. DOI: 10.1177/1474904116681798.
- Tröhler, D. (2014) 'Tradition oder Zukunft? 50 Jahre Deutsche Gesellschaft für Erziehungswissenschaft aus bildungshistorischer Sicht', *Zeitschrift Für Pädagogik*, 60/1: 9–31.
- Tuunainen, J. (2005) 'Hybrid Practices? Contributions to the Debate on the Mutation of Science and University', *Higher Education*, 50/2: 275–98.
- Waldow, F. (2009) 'What PISA Did and Did Not Do: Germany after the "PISA-Shock"', *European Educational Research Journal*, 8/3: 476–83.
- Weingart, P. (1997) 'From "Finalization" to "Mode 2": Old Wine in New Bottles?', *Social Science Information*, 36/4: 591–613.
- Weishaupt, H. and Rittberger, M. (2012) *Bildungsforschung in Deutschland: eine Situationsanalyse*. Frankfurt am Main.
- Whitley, R. (2011) 'Changing Governance and Authority Relations in the Public Sciences', *Minerva*, 49: 359–85.
- Wissenschaftsrat (WR) (2001) Empfehlungen zur künftigen Struktur der Lehrerbildung. <<http://www.wissenschaftsrat.de/download/archiv/5065-01.pdf>> accessed 13 Dec 2015
- Zapp, M., and Powell, J. J. W. (2016) 'How to Construct an Organizational Field: Empirical Educational Research in Germany, 1995–2015', *European Educational Research Journal*, 15: 537–557.
- Ziman, J. (2000) *Real Science: What It Is, and What It Means*. Cambridge: Cambridge University Press.