BRITISH JOURNAL OF SOCIOLOGY OF EDUCATION
2025, VOL. 46, NO. 3, 323–340
https://doi.org/10.1080/01425692.2025.2454316

Before (e)valuating: student testing in History and Engineering

Herbert Kalthoff and Fabian Koelsch

Institute of Sociology, Johannes Gutenberg University, Mainz, Germany

ABSTRACT

University examinations categorise students according to their individual achievements determined by teaching staff. This procedure serves the elicitation and certification of student knowledge and thus reproduces academic hierarchies. Drawing on empirical evidence from ethnographic fieldwork in Engineering and History departments, this article investigates the processes involved in designing and administering higher education examinations. It analyses the reciprocal relationship between lectures and examinations, the standardisation of lecture content through exam questions, and the use of administrative documents as examination infrastructure. The university examination is conceptualised as a distributed activity, involving various university units, each with its own specific logic, yet whose functions converge and overlap within teaching staff. The article argues that the growing significance of examinations, driven by national and global higher education reforms, is reshaping academic teaching practices in profound ways.

ARTICLE HISTORY
Received 29 May 2024
Accepted 10 January 2025

KEYWORDS
university examinations; teaching and assessment; qualitative higher education research; infrastructure of education; Germany

1. Introduction

For many years, research has investigated the testing and evaluation of material things and the context of the procedures in which this testing and evaluation are organised (see Heuts and Mol 2013; Meyer and Wilbanks 2020; Marres and Stark 2020).[1] At the same time, the testing and evaluation of human beings have increasingly become the subject of research (cf. Lamont [2012] and the journal *Valuation Studies*), with the evaluation process being more prominently the focus of analysis (Filer and Pollard 2000; Musselin 2022). Building on this body of work, this article shifts the focus to an earlier phase in the valuation process: the performance of university teaching and examinations. From the perspective of the academic disciplines, teaching represents the transfer of subject-specific knowledge whose (gradual) success is assessed through continually performed oral and written examinations – in Bourdieu's (1996) sense, it is the symbolic violence of an inculcation. From the students' perspective, it is a process of adopting this knowledge, which is linked to a skilful professional practice and understanding (Goodwin 1994; Tonso 2006). Consequently, the

CONTACT Herbert Kalthoff herbert.kalthoff@uni-mainz.de

examination process with its material infrastructure, its temporal-spatial management of students, its topics and results, provides the legitimising basis for the subsequent assessment. However, this process poses significant challenges for examiners. First, they must be able to derive examination topics from their lectures – for the students and for themselves – in a justified and verifiable manner. Second, they must design tasks that can be realistically completed within the allotted time. Third, they must weigh the complexity and difficulty of the examination topics in such a way that the examination can be passed.

Our study seeks to describe and analyse how examinations are integrated into the everyday work routines of academics in Germany. The German higher education system is characterised by distinct features that set it apart from other national systems. A first feature of the education system is its federal structure: universities fall under the jurisdiction of states (*Bundesländer*) rather than the German federal government.[2] They are responsible for decisions regarding the establishment, organisation, and funding of universities. This leads to the second feature: apart from a comparatively small tuition fee that students pay each semester, studying is free of charge. A third key feature is the autonomy afforded to professorial chairs – the central administrative units led by professors. This autonomy is guaranteed by Article 5, Paragraph 3 of the German constitution (the 'Grundgesetz' of 1949): 'Arts and sciences, research and teaching shall be free. The freedom of teaching shall not release any person from allegiance to the constitution.'[3] This implies that professors are basically free to choose the topics they teach within the scope of their denomination (e.g. 'Sociology of Education' or 'Sociological Theory'): no dean, no university president, and no minister of a federal state can compel them to cover specific topics. In practice, this means that they have significant flexibility in choosing their teaching and examination methods, allowing for a range of instructional and assessment styles. This flexibility also leads to practical challenges concerning teaching and exam design. Although there is significant autonomy, the timing of teaching and examinations is structured and framed by the universities' organisation. During the semester, the focus lies primarily on teaching activities, culminating in examinations at the end of the term, which can take various forms, including written work, oral evaluations, or tangible exhibits. Furthermore, the last 20 years have seen a marked increase in the volume and intensity of examinations at German universities. Historically, studying in Germany involved one intermediate examination after the foundation course, and students typically earned their degree after nine semesters by passing a single major final examination. As one major result of reform processes of German universities (see Fuehr and Furck 1998; Dienel 2024), examinations have become central organisational activities for academic staff. This applies to both oral and written examinations in general, but particularly to large lectures and their examinations. Since the 1960s, the 'opening' of West German universities to a more diverse student population from various social classes has led to a massive expansion of universities and degree programmes as well as to significantly larger student cohorts. To adequately manage these large numbers of students, successive selection mechanisms were established in the early phase of studies. The written examination was and is a central pragmatic tool for this early selection process. In our study, we therefore concentrate on written examinations, although other forms of assessment (e.g. assignments) also occur in the degree programmes examined. There are two reasons for this focus: first, taking written examinations is a central dimension of study in both subjects, and the analytical findings reflect the situation of these two subjects (History; Engineering). Second, examinations lend themselves to observing the interaction

between the administration and the academic sphere of a university. Written examinations are therefore particularly suitable for providing conceptual answers to the empirically undervalued question of how university teachers prepare and conduct their examinations and which forms of administration (documents, lists, etc.) they use.

In the following, we will first outline the state of research and our theoretical considerations (Section 2). We then describe our methodological approach as well as our empirical research (Section 3). In the analysis section, we then examine how the content of university teaching is translated into exam questions (and vice versa) and how the examination process itself is organised. In History, we analyse how examinations are anticipated in teaching (Section 4.1) and how small organisational forms structure the examination (Section 4.2). In Engineering, we investigate how teaching is aligned with examinations, how the standardisation of knowledge is worked upon (Section 5.1), and how examinations are conducted (Section 5.2). Finally, we contrast our empirical findings, outline the role of university examinations, and sum up our results (Section 6).

2. State of research: sociology of higher education

While academic research has been the subject of sociological studies of science for many years (see the classic Latour and Woolgar 1979; Knorr-Cetina 1981), universities as organisations of teaching and examination have hardly been investigated sociologically to date (but see Holstein 1983; Decker and Hoffmann 2023). Scholars discuss universities either as organisations within the science system (Frickel and Moore 2006) or as social sites whose educational mission is framed by a (trans-)local governance (Espeland and Sauder 2016). Concerning social selection of university education, at least three perspectives can be recognised: Bourdieu (1996) argues for a cultural reproduction of socio-economic power through elite universities, Brand and Xie (2010) show that students who are unlikely to attend university will benefit significantly from it (negative selection hypothesis), and for Luhmann (2002) universities fulfil the function of organising professional careers, which is why he understands examinations as instruments for making transformations visible and establishing legitimate performance inequalities.

Furthermore, research has focused more intensively on university studies, and their conditions: the studies show how universities compete for (very) good students, who in turn influence the performance of their fellow students (Stevens 2007), they examine the success or failure of students (Campbell 2015), and the socio-structurally shaped interactions of students with their professors (Jack 2016). Critical voices claim that structural changes towards a neoliberalisation of universities are also reflected in teaching and assessment practices, for example in an observed standardisation and formalisation of examinations (Hursh and Wall 2011). Evidence is provided as to how teaching staff cope with large courses by standardising teaching and examination materials (Schimank 1995).

Within the sociology of higher education, little is known about the practices of (student) assessment in academic subjects (Kingston 2001, 92).[4] Thus, recent research has dealt with this topic both empirically and theoretically (e.g. Jedemark and Londos 2021). A significant portion of these studies follow a measurement theory: In their perspective, university examinations generate objective and valid results, and thus accomplish a one-to-one representation of the students' performance (Ioannidou 1997; Von Stumm, Hell, and Chamorro-Premuzic 2011; for an overview, see Nieminen, Bearman, and Tai 2023). An essential part of this

theoretical and methodological approach is the assumption that the evaluators are neutral instances of a process of institutionalised human categorisation. Other studies challenge this perspective by focusing more on the practices of assessment, valuation, and evaluation, as well as the involvement of evaluators, highlighting the contingent nature of the outcomes of these processes. For instance, it is demonstrated how (everyday) theories shape local teaching and assessment profiles (Stratton and Mannix 2005), how grading is influenced by subject-specific evaluation cultures (Tsarouha 2019) and by teachers' background knowledge of students (Holstein 1983), and how academic evaluation affects students' self-image (Nieminen 2021; Nieminen and Yang 2024). Jedemark and Londos (2021) show that university faculty, through their specific use of standardised testing technologies (e.g. multiple-choice [MC] examinations), respond to institutional demands for greater transparency in academic judgements, albeit in varying ways. At the same time, MC examinations influence how students memorise key course content (Greving, Lenhard, and Richter 2023), referring to the standardisation of curricula and addressing how students prepare for examinations using past test papers (Tyagunova and Greiffenhagen, 2021). Moreover, it is claimed that university examinations are not only to be understood as situational events but are embedded in social discourses. Analysis shows how university teaching staff orientate their evaluations to discourses in which economic, political, and scientific actors articulate expectations regarding the grade value (Forde-Leaves, Walton, and Tann 2023).

Existing analyses of academic teaching and assessment in higher education have largely overlooked the activities involved in the preparation and administration of examinations – activities that are essential to the infrastructural conditions for evaluating and certifying students' performance. This article addresses this gap by examining the knowledge of those who shape university teaching and examinations. Adopting a practice-theoretical approach (e.g. Schatzki 2002; Bourdieu 1996), the article conceptualises assessments as a social mechanism that produces tested knowledge and makes students examinable and comparable. For our contribution, three questions are of interest: How does teaching anticipate and organise the construction of examination tasks? How do examinations pre-structure what will later be recognised as exam results? And how do teaching staff reflect on and observe their own practices while designing exam questions?[5] From a sociology of time perspective, teaching organises an anticipation of the assessment challenging the idea of a linear organisation of university teaching and assessment (teaching – learning – assessment – evaluation). Instead, a more circular approach is proposed, where future events (e.g. examinations) are embedded in present practices (e.g. teaching) (cf. Breidenstein 2023).

It is further assumed that assessment at universities cannot be reduced to a singular event (e.g. the examination itself) but must be understood as an activity or practices distributed over time, involving various organisational units (teaching, administration, examination boards, legal department, room administration, technical support) (see e.g. Enfield and Kockelman 2017). For instance, preparatory measures taken by these units frame the examination before it begins. During the examination, the administration is represented through official announcements and documents. Therefore, teaching staff simultaneously embody administration (list management), legal department (instruction and supervision), room administration (parcelling), and technical support. This leaves room for interpretation with regards to the binding nature of rules and specifications, which this article will explicate using the example of the list.

3. Methodological remarks

The empirical foundation of this article comprises participant observations, interviews, and the collection of field-specific documents related to the teaching of History and Engineering at German universities. With History and Engineering, this article contrasts two university subjects that perform and design their teaching differently. History, characterised by its internal differentiation across geography and epochs, is understood as a more academic discipline. While it is less governed by a strict hierarchical curriculum, as is known for German schools, it still requires knowledge (such as foreign and old languages) and is oriented towards specific written materials (archives, sources). The methodical documentation and (hermeneutic) analysis are taught separately. In contrast, Engineering is considered more structured and application-oriented. The aim of this very heterogeneous subject, which integrates many courses of study (including e.g. production engineering, computer science, architecture, urban planning), is above all to impart and develop technical and applied knowledge. Core competencies include mathematical, scientific, technical, and economic expertise, which students must acquire and apply. Qualification theses are dedicated, for example, to production processes, modelling of processes, questions of navigation, and planning of urban quarters; here, in particular, technical objects are created, but also planning drafts, which are the subject of evaluation. In summary, History represents a more traditional discipline characterised by adherence to classical educational and academic ideals, while Engineering is more market-oriented than academically focused.

The empirical data include observations in Engineering and History, and contrasts, therefore, two different epistemic cultures. Theoretical sampling began with an examination of the internal differentiation of the subjects studied (e.g. 'Ancient History', 'Early Modern History', and 'Modern History' in the case of History; 'Urban Planning' and 'Hydraulic Engineering' in the case of Engineering). Regarding Engineering, we did field research at four (technical) universities across different domains (e.g. construction, environment, mobility, computer science, urban planning, process engineering). We participated in 15 courses (lectures, project seminars, exercises, laboratory projects) and 26 examinations (written and oral presentations) and collected documents. In addition, we conducted nine guided interviews and 21 ethnographic interviews with lecturers on exam design, assessment, and (digital) teaching. In History, we conducted six weeks of fieldwork in an institute of History: 10 courses (lectures, seminars, tutorials, exercises, colloquia, etc.) and three written examinations were observed. This was followed by brief ethnographic stays at two universities, where we observed examinations, valuation practices, and collective evaluation procedures. Within the context of our ethnographic research, we conducted 31 ethnographic interviews with academic staff (professors and assistants). Additionally, we conducted 11 guided interviews with historians (professors as well as assistants) in eight German states at nine different universities, across different domains (e.g. early modern History, modern and contemporary History, non-European History).

Guided interviews (in German: *Leitfadeninterviews*) are understood as interviews conducted by social scientists themselves using a previously created guideline; this represents a more formal, traditional interview situation. Such a guideline consists of a list of topics that are addressed in the interview, although the interviewee's statements must be continuously checked in terms of their relevance to the research interest; this includes potential deviation from the guideline (cf. Hopf 2016; Fontana and Frey 2000). We conducted ethnographic

interviews, understood as 'friendly conversation' (Spradley 1979, 55), before and after classes or examinations, in the participants' offices, in the hallway during a coffee break, or in administrative offices. Ethnographic interviews therefore leverage the researcher's mobility and allow addressing their actual observations to the participants (cf. Spradley 1979). The interviews were transcribed, ethnographic protocols revised, and open-coded using the principles of Grounded Theory (Strauss 1987; Emerson, Fretz, and Shaw 1995). Collected documents were also analysed in terms of their design and impact (Atkinson and Coffey 2011). We understand open coding as a process of translating empirical observations into analytical categories. Through this practice of an intensive fine-grained analysis, familiarity with the subject is overcome, and the sociological explication is freed from the participants' practical concerns.

4. History – between daily business and discomfort

The study of History is characterised by an orientation towards the processing, contextualisation, and critical analysis of historical sources. In view of the great diversity of source material, History teaching covers a wide range of topics: from analysing historical video recordings, photographs, or written testimonies, to practising cultural skills and visiting memorial sites, museums, and archives. Regarding the assessment of students' knowledge, a tension emerges between a more traditional understanding of education and standardised examination management.[6]

4.1. The anticipation of examinations in lectures

Despite recurring contention, it is hard to imagine university teaching – whether in History or other disciplines – without lectures. As a characteristic genre of speech, they have been analysed either in terms of their asymmetrical structure (Friesen 2017), as an institutionally granted authority to speak to an audience, or as a form of collective thinking (Masschelein and Simons 2013). Most studies conceive of lectures as a purely educational format, decoupled from examination processes; as already outlined, we argue that they are already part of it (e.g. Breidenstein 2023). When asking how academics coordinate teaching and examination programmes, various aspects of mutual configuration come to the fore. Peter, a lecturer, commented on the use of the lecture manuscript as follows:

> I have the finalised manuscript with me as a fallback option, so that I know what I have presented. This issue is important to me because it is subject to exams. In the end, I have to be sure that what is part of the exam has been part of what I told.

Here, the lecture manuscript acts as a temporal reassurance, so there can be a fit between oral speech and final exam questions. It is evident that the lecturer only considers what he has articulated, while leaving unacknowledged what may have been heard by students; a statement common in this academic field. By that, teaching staff adhere to the fiction that all students – regardless of their social position – interpret and internalise the content uniformly (Jack 2016). Such an everyday theory of university teaching simultaneously stabilises one's own role as a lecturer and serves as a prerequisite for attributing individual performance. Furthermore, the manuscript is not only the memory of the lecture and a template for the examination, but also a guideline of the lecture itself, which does not – in this case – allow deviation. Conversely, examinations can shape lectures, particularly when

the manuscript is designed with 'good' exam questions in mind. Due to the autonomy of teaching, it comes to varying styles in this regard: while most of the professors observed rely on their manuscripts and read them verbatim, others deliver lectures that are more spontaneous. Both forms nevertheless integrate dialogical elements into their lecture (e.g. student questions at the end).

Generally, and not surprisingly, parts of the lecture become examination knowledge, transforming ephemeral speech acts into materialised writing and abstract examination tasks. In the following example, content is interactively elicited and related to the examination:

> Frederic, the lecturer, says: 'I would like to take a look at the bull […] Let's start with the first [question]: who?' A student from the front row answers briefly: 'Originator is Mescal'. Frederic confirms and continues: 'To whom?' Again, an answer from up front: 'Addressed to Marco Pablos. Ruler of Castille'. Frederic comments: 'Very good! Can you also tell us the life dates?'. The student replies, 'Ehm (2) No'. Frederic turns to the crowd: 'Okay. Once you describe the source in the exam, you ought to do so'.

This excerpt highlights a routine approach in teaching History: attention is centred on a historical source text and then interrogated with regards to its conditions of origin. Here, the pragmatic hint or advice regarding the life dates is important. The lecturer states that these kinds of dates are not only to be known, but also to be stated in the examination: an 'ought' that is a 'must'. Through this, contents of the lecture are related to the examination in speech acts of the lecturer: it is a modulation of simultaneous presence and absence. Later, the same historical example appears as an exam question:

> The Dei Veritas bull of 1181 plays a central role in Castilian supremacy. Get to the heart of the meaning of the bull through listing central aspects (6P) …

In its brevity, the formulation of the exam question abstracts from the previous lecture session and yet calls upon it again by naming the bull. It is striking that the way in which the question is formulated reduces the complexity of the case. There is no open question about the meaning of the bull; rather, its meaning is presented as a fact – provided with contextual data (year, region, etc.) – and thus it induces students' answers. It also shows how lecturers are involved in the production of exam results through their questions. Notably, details like life dates mentioned in the lecture are omitted in the examination task. Concerning historical knowledge that should be taught and tested, our empirical observations reveal a certain tension regarding the normative standards: on the one hand, actors advocate a stronger cultural-theoretical orientation of the subject and relativise factual knowledge ('against historical positivism and factology' [Peter]). On the other hand, there is the idea that historical facts are important for the subject, to be dealt with by teaching staff and to be learnt by students. Maria, a professor, stated: 'And if it's also stupid memorisation of data series, then you just have to learn stupidly by heart, right?'. This tension is reflected in the design of the examination: from open essay-oriented questions to MC examinations as the *ultima ratio* of a mass study programme. With respect to MC examinations, Frederic stated:

> […] First of all, you think, gosh, as a historian, multiple-choice, how dare you! But with regard to a hundred students, you turn pragmatic, and I had a great role model. In the past, (lectures) didn't have to be tested anyway but with Bologna they started to be. Then, my former professor started designing MC questions and I found these quite intelligent.

The importance of MC questions lies in managing large examinations in an available worktime. It thus stands for a social ageing within the university ('turning pragmatic'). While the time required preparing questions increases, it is significantly reduced for correction through formalisation. This addresses a core principle of modern examination organisation: standardisation and rationalisation. According to the historian, its form does not determine its content – despite its standardising effects – as far as 'intelligent' MC questions are feasible. It is this contradictory position – being squeezed between pedagogical and bureaucratic form – to which the transformation of the lecture material into MC questions runs. According to Bowker and Star (1999), standardisation is also a process of establishing uniformity. Firstly, the consistency in form does not only represent the fiction of equality, it also co-produces it. Secondly, standards mediate between different fields and enable cooperation and division of labour (e.g. between teaching staff and the examination office). Other historians in our interviews also addressed the distinction between curriculum content and standardised examination format. Gregor, one historian, stated:

> The exam is taken on a PC and the system corrects it automatically. Multiple choice, that's clear. And then my assistants check the fill-in-the-blank texts first. They are corrected again by hand to see, okay, there's someone who only has, say, 32 percent of the points and then you ask: Can you find anything to give a point […]? And then I look at it again, especially about the doubtful cases.

The assessment of an MC examination by 'the system' is possible because the subject knowledge has previously been formalised in standardised question–answer formats. For the calculation of correct answers ('points'), everything that can be formalised can also be mechanised (Heintz 2003). Against this mechanisation, which is carried out by the software and which can be described as a numerical form of objectification, the historian brings into play a social form of objectification, the aim of which is to subsequently ensure that all assessment improvements have been exhausted. After the numerical objectification, he provides a two-stage procedure: research assistants go through all computer-based assessments again and, in critical individual cases, the professor, with his authority, decides. In essence, this means that answers that are professionally unclear or almost wrong can be recognised as answers for which points can be awarded. In this act, the standard is shifted: it is not the technical correctness alone that counts, but the recognition of the answer as a correct one, which in turn enhances exam results. With this re-assessment of the technically produced results, the historians observed secure their professional authority and sovereignty: the assessment of the 'system' becomes their assessment of the examination.

Exam questions in History, which do not follow the MC mode, do not only test students' knowledge, but also train them to use the academic vocabulary of the subject. It is an exercise in the cultural technique of historical reasoning: arguing for or against a controversial thesis, describing and interpreting sources or critically discussing key concepts of the subject of History. The exam questions can therefore also evoke the idea of a research community with which the candidates must engage. For historians who resort to the MC examination, a pragmatic attitude is about managing their time; for those critical of MC, the MC examination symbolises a measurement-theoretical approach to education, which is foreign to the subject and therefore undermines their humanistic ideal of academic education. In other words, it is obvious that the standardisation of curricula is not solely the effect of a new governance of universities (Hursh and Wall

2011), but is understood as a procedure with which teaching staff pursue their own goals: an adaptation to the changed conditions and the preservation of the ideal of university training.

4.2. Audit 1: working (on) the list

The conduct of university examinations depends on various resources and materials, such as attendance lists, examination instructions, or administration programmes. In contrast to the previous observations, these artefacts are not merely pedagogical, but bureaucratic forms. Despite their ubiquity and simplicity, lists are an effective artefact providing insight into the organisation of examinations. Different studies (e.g. Goody 1977; Stäheli 2012; Fludernik 2016) have emphasised the ordering role of lists: they combine and arrange things (e.g. enrolment numbers, signatures), produce clarity and comparability, and make differences recognisable. When following the handling and circulation of lists, one can observe how different instances and players are involved in the organisation of university examinations and assessments. For example, the attendance list does not guarantee how many students will take the examination. In the case of a History examination, it functions as a bureaucratic artefact that mediates between planning and implementation. During the examination, the list made visible who appeared and who did not. It is not a closed list, but can be expanded; for example, students who bring a legitimate document can be added to the list. In these cases, the examination office acts as a supervisory body. While the list is a document that may be manipulated by the examiners, it is a sign of a powerful, bureaucratic principle of order to which the examinees must submit. The following is an excerpt from the observation log:

> A student places his completed exam sheets on the submission pile, takes the attendance list […] To Bruce, who rushes over, he says: 'I'm not on the list'. After short consideration, Bruce asks whether the student has registered. At first, the student replies 'yes', but then adds that he is no longer sure. Bruce calmly replies that a solution will certainly be found, and that he should write his name and signature on the back of the list to be safe […]

This interaction highlights the power of the list, specifically regarding who is legitimately taking part in the examination. Not being on the list is tantamount to a loss of status; that is, being delegitimised as an examinee. The student's uncertainty about their status at the end of the examination is met by the sovereignty of the lecturer in dealing with the list. It is not sacrosanct, but a pragmatic means of documentation, in which digitally generated and handwritten entries complement each other. This act of supplementing the list is a workaround, not a full resolution of the issue. Instead, the addition delays the problem for later verification, as all new entries are subsequently cross-checked (Schabacher 2022). As has been shown, the list can and must become an object of writing: its function as an attendance list is only fully realised through the students' signature. The relevance of this process lies in the specificity of this act of writing: the students do not write anything, but they sign with their name and in their handwriting. Thus, by signing the list, students testify to their presence as examinees as well as to the sovereignty of their writings to be kept for further processing during the timespan of the list. Like other examination objects, the list thus institutes the singularisation of the examinees and 'their' knowledge.

5. Engineering – between pragmatism and standardisation

Engineering studies in Germany are characterised by large lectures with an anonymous audience, which mainly take place in the introductory phase. Their features are high numbers of participants (between 80 and 500), their significance in the curriculum (ECTS credits), and the duration of standardised examinations (between two and four hours). For instance, participants referred to an examination for the introductory lecture in electrical engineering as 'the Great Arithmetic'. On the part of the examination offices and dean's offices, this examination activity implies monitoring, in which the performance level of students is systematically recorded and checked on at an organisational level.[7]

5.1. Teaching and the design of the examination

Lectures in the Engineering curriculum attract varying degrees of participation depending on their design and orientation. While some academics downplay the absence of students ('they do it on their own'), others try to meet and react to this challenge. For example, Christian, a professor, works towards narratively embedding 'dry' foundations in identificatory tales:

> In hydraulic engineering, stilling basin design will be on the agenda tomorrow. There, I can just show them an example of the Oroville Dam in California, where the stilting didn't work, when the dam started. 180,000 people had to be evacuated because it was no longer stable. So, they see what can happen and why a learning objective is being taught […]

While the evocation of technical failures and accidents is intended to make the course content plausible, relevant, imaginable, and experienceable (see also Wylie 2019), the programme of the lecture (stilling basin design) refers to arithmetic formulas that are taught in the session and then practised in exercises accompanying the lecture. Accompanying the lecture, exercises are explicit teaching formats for the training of complex arithmetic and computational procedures. Traditionally, Engineering is characterised by a dense schedule of examinations and an affinity for measuring, testing, and certification. Depending on the specialisation, the approach varies: while in design-oriented fields such as architecture, urban planning, or industrial design, expert review in the form of professional criticism plays an important role, in technology-oriented fields, review and control are linked more tightly to instrumental and computational methods. Thus, the two forms of objectification (see earlier) also characterise the examination process in Engineering combining writing and orality, calculation, and expression. However, they also open up the scope for correction after correction; that is, for adjusting the exam results.

For Engineering, the overall question is whether an examination proves to be examinable and thus feasible, which only becomes apparent after the examination has been marked. Thus, the exam results are not only related to the (non-)knowledge of students, but also to the difficulty of the exam questions teaching staff have constructed. Leona, a professor, stated:

> […] if 100 people write an exam and there is not one who would have at least a 1.9 or 1.7, we go through our own work again and ask: Was it really more difficult this time? Which tasks were the problem? And then […] we adjust the grade level. But if there are a few people who achieve straight 1.0's, then I think it was attainable. And if […] everyone else is at 3.7, then we don't have to revise anything.

Leona clarifies how the assigned grades provide information about whether the exam questions she developed were adequate in the sense of corresponding to the students' knowledge so that a distribution of grades is achieved that can be read as a normal distribution.[8] Thus, the assessment of the examination allows for lecturers to review their own construction work, which was associated with ideas about what the students should know. The results also answer the question of whether there is a need for action ('adjust the grade level'). In this framing model, bad grades are not attributed exclusively to the students, but also reflexively to academics' own exam design. In this case, it is mainly the 1.0 results that inform whether the examination was feasible. However, this perspective – 'adequate' exam questions result from 'correct' exam results – leads into a circle that cannot be closed.[9] Again, a second anticipation: the professor thoughtfully runs through different scenarios that could imply a change in the correction of the examination. This means, the work on the interplay between exam design and exam results highlights how the actual practice of evaluation challenges the theoretical modelling of evaluation as an authentic representation of performance (Nieminen, Bearman, and Tai 2023).

In contrast to courses, the written examination interrupts the educational and learning principle of the university and switches to the assessment of the knowledge documented in the examination. The central means of assessment are the exam questions. Open questions, for example, are often associated with the assumption that students (can) show their independent thinking, while closed questions generate a better comparability of results. Therefore, exam questions communicate different requirements and expectations of students, which are subject to individual and subject-specific preferences. Rahim, a professor, stated:

> For the large lecture, few parts are multiple-choice questions, but the rest are arithmetic tasks. […] You can imagine it this way: there is a technical problem, and it is not identical to the tasks in the exercises. […] I'll give you an example: in the exercise, a generator for a wind machine was discussed. Well, the same principle can be used for a car, for an engine. And that's what they get in the exam.

Rahim reports a linearity of teaching format, exercises, and written examinations. Written examinations are based on the exercises, which are based on the lectures. This means that what was tested in the exercises is reorganised with other data and materials in the written examinations. With the reorganised examination tasks, the participants rely on the fact that the students are familiar with the knowledge through repeated exercises. Over the years, individual questions vary, new ones are added, and old ones are removed, but can be reintegrated later. Jens, a research assistant, commented on this: 'If you do it skilfully, you practically prepare the new task first as an exam question so that you can use this task in the next year's exercises'. This economy of task design switches back and forth between the novelty value of tasks and their reuse in the examination or teaching formats, so that tasks can circulate. Thus, exercises function as an intermediary between lecture and examination. Participants also report on how they design examinations with a mixture of reproduction questions (usually asked as MC questions) and the aforementioned reorganisation questions ('calculation tasks'). Transfer questions (in the strict sense) are regularly omitted; this lowers the level of difficulty of the tasks.

Such an examination may be made up of three major tasks, each of which is structured as follows: task title, task text, (technical) illustration, further information, and technical

data. For example: 'Draw magnetic field lines and the equivalent circuit diagram' or 'Calculate the number of turns N1 of a stator winding, considering the copper current density'. In a separate line, points are given for each task, which assign a value to the tasks to be completed. The formal structure of the examination and the structure of the tasks clearly show how the students' activities are expected and standardised: reading, calculating, and drawing. During the sequence of tasks, the technical drawing is converted into a calculation and the calculation results in a value that is to be drawn on a graphical diagram. With these translations, the task stands for a chain of different mathematical and diagrammatic representations. Therefore, visual abstraction and the concrete design of circuits are systematically tested in the examination. At the same time, students are addressed by the examination tasks as prospective engineers who must solve these tasks. Other cases are also possible: for instance, how the design of the tasks is guided by the consideration that not all tasks can be completed in the given time of the examination. In this way, the examination functions as an initiation ritual, which regulates the threshold of passing upwards. In this situation, the students prove that they belong to the profession 'simply' by passing, with less emphasis on the specific grade achieved. The design of written examinations in Engineering reflects an epistemological commitment to measurement and calculation. It allows the undisturbed circulation of task and question types between courses and examinations, turns every (practical) exercise into preparation for the examination, and measures success based on normally distributed results. For its part, the normal distribution stands for the participants' belief in an objective measurement of performance that is independent of social disturbance factors.

5.2. Audit 2: lists and physical checks

Carrying out an examination also means administrative work for teaching staff. The variable number of registered participants in advance of the examination is relevant for organising capacities, copies of exam sheets, and the number of supervisors. On the day of the examination, the comparison of the list of participants with the physically present students is the first formal act. This process addresses three key concerns: access (who is allowed?), distribution (how many per room?), and permitted objects (what may be used?). An excerpt from an observation log illustrates this process:

> The seats are marked with seat numbers and stand in great distances from each other. On each of the tables lies a sealed envelope with the exam sheets. […] A staff member addresses the students: 'Good, let's start by clarifying attendance. I'm going through the list now, please tell me your seat number when answering whether you're here'. Name after name is read aloud and the students answer with their seat number.

The meticulous documentation of the seat numbers produces an index that links the list of participants with a seating plan, delineating the spatial order. By equipping the room with clearly separated seats, not only are the examinees separated from each other (Foucault 1977), but this spatial configuration partially shifts control from supervisors to the physical layout of the room itself. It becomes clear how the documentation is structured by paper files (slips of paper with seat numbers, seating plan, lists), intervening into space, marking positions as (not) legitimate, and reducing the choice of seats. In addition, the address transforms the encounter into a formal situation: the reading out of names breaks with the

anonymity that is characteristic of university teaching, but simultaneously evokes a fundamental unfamiliarity with the persons present. The supervisor continues the briefing by reading aloud some general notes:

> 'Only those exam aids that are expressly permitted may be used. […]'. If we find out later that you are wearing something like that, it's bad […] So [reads again]: 'Anyone who contravenes or violates this through cheating will be excluded from further participation in the examination. […] Anyone who withdraws after the exam has started can expect that the exam will be declared failed' […] It's seven minutes past eight o'clock. We are guiding ourselves […] through the clock here in front.

This excerpt documents a typical oral recitation of information available in writing, enforcing the rules for this event – their authoritativeness is increased through further sanctions. Once attendance is checked, it will be double-checked. This validation finalises the admission procedure: anyone who is present, registered (list of participants), sits on a numbered chair (seating plan), and follows the instructions can take the examination. From now on, certain behaviours and aids are considered cheating. In this context, the documentation of seat numbers also becomes relevant: if a strong similarity of wrong results or drawings is noticed during the correction, a cheating attempt can also be claimed retrospectively. The available time is also generalised: only the time on the clock visible to all is valid. Spatial distribution and time synchronisation conclude the formal start of the examination.

After the examination has started, one supervisor walks through the rows and checks ID cards and utensils, the other casts his gaze into the room. The ID check underlines the official character of the situation, and the students testify to their presence with their signature. This fourfold verification – list of participants, seating plan, ID check, signature – exemplifies a stringent control regime in Engineering. Its purpose is to uphold the fiction of individual performance, ensuring that each examination unequivocally corresponds to a specific student (Meadmore 1993). The supervisor represents the codified guidelines of the examination organisation in whose name they speak. The guidelines are also materialised in the infrastructure of the overall setting. Thus, the telos of this practice lies at least partly in the creation of conditions that enable objective judgements about student performance.

6. Discussion and conclusion

Using the example of two academic subjects – History and Engineering – this article described and analysed how university examinations in Germany are materially organised and prepared in terms of content. In this context, the article argued that the examination process should be understood as an organisationally 'distributed activity' (Enfield and Kockelman 2017). By doing so, the article addressed the observation that a sociology of evaluation focused solely on situational acts of evaluation risks rendering trans-situational elements invisible, thereby 'humanising' and individualising evaluation (cf. Holstein 1983). That is to say that the sociology of (e)valuating is faced with the challenge of dialectically relating the situational accomplishment of testing and grading students at universities, on the one hand, and the structural conditionality of this constellation and practice on the other (Waibel, Peetz, and Meier 2021). In the empirical part, the article examined cultural

forms and material infrastructures through which the university organises the examination process. The organisational level is not an autonomous domain detached from teaching and research. Instead, it was shown how academic teaching staff participate in organisational practices, and how knowledge circulates between teaching and examination practices: constructing exam questions, auditing and monitoring examinees, and so on.

In Engineering, the way in which examination is performed mirrors exactness and precision – values deeply anchored in the self-image of the discipline. While in teaching, the examination is prepared through exercises, the examination and its local standards are consolidated through the monitoring of the exam results. Control mechanisms do not stand for something external or foreign to the subject; rather, they secure the acceptance of individual achievements by enforcing equal examination conditions, stabilise the social situation as self-sustaining, and communicate a shared understanding of what constitutes a valid test situation. In History, examinations are more contested. They are regarded as an appendage of large mass lectures and stand for the demands of a modern examination organisation, which cannot always be reconciled with the educational theories of historians. Here, an internal contrast between affirmation and critique of the various forms of examination emerges, which provides for different professional self-images. These differences manifest themselves within the subject as positional struggles between those who want to maintain certain standards and those who want to establish a new image of the discipline.

Regarding the temporal dimension of assessment and evaluation, different forms of relating teaching and examinations occurred. In History, the first form is about a present recourse to the past (the lecture) to design the near future (the examination). The second form is based on an anticipation of the future (the examination) in order to orient the present practice of the lecture towards it. In contrast, Engineering organises a temporal circularity of subjects with a high degree of standardisation of teaching: exam or exercise questions that the lecture deals with or is based on are later fed back into the lecture–exercise–exam cycle as recycled questions. The exercise acts as the central temporal hinge that mediates between the lecture and the examination.

In both academic subjects, the social is actively mobilised in teaching, and is individualised by the conditions of the examination, so that, later, it can be assessed and compared as personal knowledge (e.g. Meadmore 1993). This modulation of the situation is made possible by various artefacts (lists, exam sheets) and activities (monitoring). Students are not only checked for their attendance, but, above all, whether they submit to the assessment of their knowledge, make themselves the object of the examination, and allow themselves to be evaluated (e.g. Nieminen and Yang 2024, 1032). They are asked to accept that they have been objectively assessed (factuality), that it is they who have been assessed (attribution) and that they are what they have been assessed as (identification).

To summarise: this article did not examine the assessment of student work, but rather the upstream, organisational, and material embedding of university examinations as well as the involvement of teaching staff. The article demonstrated that examination events are enacted via a distribution of different activities, each following their own logic (of teaching, ordering, controlling) revealing a complex – and not always frictionless – adaptation to educational policy ideas: the comparability of results, the standardisation of exam conditions, and the permanent institutionalisation of examinations every semester. Although the article reflects on the conditions of German universities, it may serve to stimulate contrasting studies on exam practices in other national systems of higher education as well.

Notes

1. This article uses 'assessment' to refer to formal situations of testing (e.g. a written examination), 'valuation' is used when an answer or an examination is marked, and 'evaluation' when the article refers to grading an examination or an institution (cf. Wiggins 1998; Lamont 2012).
2. German universities differ according to their legal form: most German universities are state institutions that are corporations under public law; private universities are subject to private law.
3. https://www.gesetze-im-internet.de/englisch_gg/englisch_gg.html#p0035. The background to this formulation is the instrumentalisation of universities and research for the purposes of National Socialism in Germany (1933–1945; e.g. Merton 1938; Macrakis 1993).
4. Issues of how students can be evaluated fairly and properly have been discussed since the modern grading system was introduced at the end of the nineteenth century (Gipps 1999). This discourse became a systematic part of the education system materialised, for example, as handbooks and manuals (Bloxham and Boyd 2007).
5. The notion of self-observation refers to the practice that teaching staff also observe their own achievements when preparing for and conducting a written examination (cf. Luhmann 2002; Kalthoff 2013).
6. This understanding of education is not easy to translate, as it refers to the German concept of 'Bildung', which could also be translated as 'formation', 'culture', or 'forming'.
7. We are very grateful to Kornelia Engert for her support in the empirical research (see Engert 2022).
8. The grading system in German universities is as follows: 'very good', 'good', 'satisfactory', 'sufficient', and 'deficient'. Grades are represented in Arabic numerals: '1' stands for 'very good', '2' stands for 'good', '3' stands for 'satisfactory', '4' stands for 'sufficient', and '5' indicates failure and is often represented as 'n.b.' (*nicht bestanden*; not passed). University teaching staff further differentiate grades using decimal numerals (e.g. a weak 'very good' as 1.3; a strong 'good' as 1.7; a strong 'sufficient' as 3.7, etc.) or they calculate the grade using the arithmetic mean (e.g. 1.9).
9. For a detailed account of this circle see Collins (1985, 79).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was supported by DFG Deutsche Forschungsgemeinschaft (German Research Foundation).

ORCID

Herbert Kalthoff http://orcid.org/0000-0003-3040-7103
Fabian Koelsch http://orcid.org/0009-0008-5893-2253

References

Atkinson, P., and A. Coffey. 2011. "Analysing Documentary Realities." In *Qualitative Research. Theory, Method and Practice*, edited by D. Silverman, 56–75. London: Sage.

Bloxham, S., and P. Boyd. 2007. *Developing Effective Assessment in Higher Education: A Practical Guide*. Maidenhead: Open University Press.

Bourdieu, P. 1996. *The State Nobility*. Stanford: Stanford University Press.

Bowker, G. C., and S. L. Star. 1999. *Sorting Things out: Classification and Its Consequences*. Cambridge: MIT press.

Brand, J. E., and Y. Xie. 2010. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75 (2): 273–302. https://doi.org/10.1177/0003122410363567.
Breidenstein, G. 2023. "Prüfungswissen." *Zeitschrift für Pädagogik* 5 (5): 568–583. https://doi.org/10.3262/ZP2305568.
Campbell, C. 2015. "The Socioeconomic Consequences of Dropping out of High School: Evidence from an Analysis of Siblings." *Social Science Research* 51: 108–118. https://doi.org/10.1016/j.ssresearch.2014.12.011.
Collins, H. 1985. *Changing Order: Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press.
Decker, K., and C. Hoffmann. 2023. "Training Scenes. Taking Science Studies to the Classroom." *Social Studies of Science* 53 (3): 402–426. https://doi.org/10.1177/03063127231164583.
Dienel, C., ed. 2024. *Globalizing Higher Education and Strengthening the European Spirit: How the Bologna Reform Has Changed Our World*. New York: Routledge.
Emerson, R. M., R. I. Fretz, and L. L. Shaw. 1995. *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press.
Enfield, N. J., and P. Kockelman. 2017. *Distributed Agency*. Oxford: University Press.
Engert, K. 2022. "Instruktion und ihre Teilnehmer: Was Ist in Einem Seminar?." In *Ethnografie der Hochschule. Zur Erforschung universitärer Praxis*, edited by D. Meyer, J. Reuter, and O. Berli, 83–106. Bielefeld: transcript.
Espeland, W. N., and M. Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. New York: Russel Sage Foundation.
Filer, A., and A. Pollard. 2000. *The Social World of Pupil Assessment. Processes and Contexts of Primary Schooling*. New York: Continuum.
Fludernik, M. 2016. "Descriptive Lists and List Descriptions." *Style* 50 (3): 309–326. https://doi.org/10.5325/style.50.3.0309.
Fontana, A., and J. H. Frey. 2000. "The Interview. From Structured Questions to Negotiated Text." In *Handbook of Qualitative Research*, edited by N. K. Denzin and Y. S. Lincoln, 2nd ed., 645–672. Thousand Oaks, CA: Sage.
Forde-Leaves, N., J. Walton, and K. Tann. 2023. "A Framework for Understanding Assessment Practice in Higher Education." *Assessment & Evaluation in Higher Education* 48 (8): 1076–1091. https://doi.org/10.1080/02602938.2023.2169659.
Foucault, M. 1977. *Discipline and Punish: The Birth of the Prison*. New York: Pantheon.
Frickel, S., and K. Moore, eds. 2006. *The New Political Sociology of Science: Institutions, Networks, and Power*. Madison: University of Wisconsin Press.
Friesen, N. 2017. *The Textbook and the Lecture: Education in the Age of New Media*. Baltimore: Johns Hopkins University Press.
Fuehr, C., and C. Furck, eds. 1998. *Handbuch deutsche Bildungsgeschichte. Band 6: 1945 bis zur Gegenwart*. München: Beck.
Gipps, C. 1999. "Socio-Cultural Aspects of Assessment." *Review of Research in Education* 24 (1): 355–392. https://doi.org/10.3102/0091732X024001355.
Goodwin, C. 1994. "Professional Vision." *American Anthropologist* 96 (3): 606–633. https://doi.org/10.1525/aa.1994.96.3.02a00100.
Goody, J. 1977. *The Domestication of the Savage Mind*. Cambridge: Cambridge University Press.
Greving, S., W. Lenhard, and T. Richter. 2023. "The Testing Effect in University Teaching: Using Multiple-Choice Testing to Promote Retention of Highly Retrievable Information." *Teaching of Psychology* 50 (4): 332–341. https://doi.org/10.1177/00986283211061204.
Heintz, B. 2003. "When Is a Proof a Proof?" *Social Studies of Science* 33 (6): 929–943. https://doi.org/10.1177/0306312703336005.
Heuts, F., and A. Mol. 2013. "What Is a Good Tomato? A Case of Valuing in Practice." *Valuation Studies* 1 (2): 125–146. https://doi.org/10.3384/vs.2001-5992.1312125.
Holstein, J. A. 1983. "Grading Practices: The Construction and Use of Background Knowledge in Evaluative Decision-Making." *Human Studies* 6 (1): 377–392. https://doi.org/10.1007/BF02127770.

Hopf, C. 2016. *Schriften zu Methodologie und Methoden qualitativer Sozialforschung*. Wiesbaden: Springer VS.
Hursh, D., and A. F. Wall. 2011. "Repoliticizing Higher Education Assessment within Neoliberal Globalization." *Policy Futures in Education* 9 (5): 560–572. https://doi.org/10.2304/pfie.2011.9.5.560.
Ioannidou, M. A. 1997. "Testing and Life-Long Learning: Open-Book and Closed-Book Examination in a University Course." *Studies in Educational Evaluation* 23 (2): 131–139.
Jack, A. A. 2016. "(No) Harm in Asking: Class, Acquired Cultural Capital, and Academic Engagement at an Elite University." *Sociology of Education* 89 (1): 1–19. https://doi.org/10.1177/0038040715614913.
Jedemark, M., and M. Londos. 2021. "Four Different Assessment Practices: How University Teachers Handle the Field of Tension between Professional Responsibility and Professional Accountability." *Higher Education* 81 (6): 1293–1309. https://doi.org/10.1007/s10734-020-00612-4.
Kalthoff, H. 2013. "Practices of Grading. An Ethnographic Study of Educational Assessment." *Ethnography and Education* 8 (1): 89–104. https://doi.org/10.1080/17457823.2013.766436.
Kingston, P. W. 2001. "The Unfulfilled Promise of Cultural Capital Theory." *Sociology of Education* 74: 88–99. https://doi.org/10.2307/2673255.
Knorr-Cetina, K. 1981. *The Manufacture of Knowledge*. Oxford: Pergamon Press.
Lamont, M. 2012. "Toward a Comparative Sociology of Valuation and Evaluation." *Annual Review of Sociology* 38 (1): 201–221. https://doi.org/10.1146/annurev-soc-070308-120022.
Latour, B., and S. Woolgar. 1979. *Laboratory Life: The Construction of Scientific Facts*. London/Beverly Hills: Sage.
Luhmann, N. 2002. *Das Erziehungssystem der Gesellschaft*. Frankfurt/Main: Suhrkamp.
Macrakis, K. 1993. *Surviving the Swastika: Scientific Research in Nazi Germany*. New York: Oxford University Press.
Marres, N., and D. Stark. 2020. "Put to the Test: For a New Sociology of Testing." *The British Journal of Sociology* 71 (3): 423–443. https://doi.org/10.1111/1468-4446.12746.
Masschelein, J., and M. Simons. 2013. "The University in the Ears of Its Students. On the Power, Architecture and Technology of University Lectures." In *Die Idee der Universität – revisited*, edited by N. Ricken, H.-C. Koller, and E. Keiner, 173–192. Wiesbaden: Springer VS.
Meadmore, D. 1993. "The Production of Individuality through Examination." *British Journal of Sociology of Education* 14 (1): 59–73. https://doi.org/10.1080/0142569930140104.
Merton, R. 1938. "Science and the Social Order." *Philosophy of Science* 5 (3): 321–337. https://doi.org/10.1086/286513.
Meyer, M., and R. Wilbanks. 2020. "Valuating Practices, Principles and Products in DIY Biology: The Case of Biological Ink and Vegan Cheese." *Valuation Studies* 7 (1): 101–122. https://doi.org/10.3384/VS.2001-5992.2020.7.1.101.
Musselin, C. 2022. "Evaluation and Merit-Based Increase in Academia: A Case Study in the First Person." *Valuation Studies* 8 (2): 73–88. https://doi.org/10.3384/VS.2001-5992.2021.8.2.73-88.
Nieminen, J. H. 2021. "Beyond Empowerment: Student Self-Assessment as a Form of Resistance." *British Journal of Sociology of Education* 42 (8): 1246–1264. https://doi.org/10.1080/01425692.2021.1993787.
Nieminen, J. H., and L. Yang. 2024. "Assessment as a Matter of Being and Becoming: Theorising Student Formation in Assessment." *Studies in Higher Education* 49 (6): 1028–1041. https://doi.org/10.1080/03075079.2023.2257740.
Nieminen, J. H., M. Bearman, and J. Tai. 2023. "How Is Theory Used in Assessment and Feedback Research? A Critical Review." *Assessment & Evaluation in Higher Education* 48 (1): 77–94. https://doi.org/10.1080/02602938.2022.2047154.
Schabacher, G. 2022. *Infrastruktur-Arbeit. Kulturtechniken und Zeitlichkeit der Erhaltung*. Berlin: Kadmos.
Schatzki, T. R. 2002. *The Site of the Social. A Philosophical Account of the Constitution of Social Life and Change*. University Park: Pennsylvania State University Press.
Schimank, U. 1995. *Hochschulforschung im Schatten der Lehre*. Frankfurt a.M.: Campus.
Spradley, J.-P. 1979. *The Ethnographic Interview*. New York: Holt, Rinehart Winston.
Stäheli, U. 2012. "Listing the Global: Dis/Connectivity beyond Representation?" *Distinktion: Scandinavian Journal of Social Theory* 13 (3): 233–246.

Stevens, M. 2007. *Creating a Class: College Admissions and the Education of Elites*. Cambridge: Harvard University Press.
Stratton, J. A., and L. H. Mannix. 2005. *Mind and Hand: The Birth of MIT*. Cambridge: MIT Press.
Strauss, A. L. 1987. *Qualitative Analysis for Social Scientists*. Cambridge: Cambridge University Press.
Tonso, K. L. 2006. "Student Engineers and Engineer Identity: Campus Engineer Identities as Figured World." *Cultural Studies of Science Education* 1 (2): 273–307. https://doi.org/10.1007/s11422-005-9009-2.
Tsarouha, E. 2019. *Prüfungspraktiken an deutschen Hochschulen*. Wiesbaden: Springer VS.
Tyagunova, T., and C. Greiffenhagen. 2021. "Learning for the Test: A Study of Students' Sharing and Discussing past Multiple Choice Questions on Facebook." *Ethnographic Studies* 18: 43–74. https://doi.org/10.5281/zenodo.5805469.
Von Stumm, S., B. Hell, and T. Chamorro-Premuzic. 2011. "The Hungry Mind: Intellectual Curiosity is the Third Pillar of Academic Performance." *Perspectives on Psychological Science* 6 (6): 574–588. https://doi.org/10.1177/1745691611421204.
Waibel, D., T. Peetz, and F. Meier. 2021. "Valuation Constellations." *Valuation Studies* 8 (1): 33–66. https://doi.org/10.3384/VS.2001-5992.2021.8.1.33-66.
Wiggins, G. P. 1998. *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.
Wylie, C. D. 2019. "Socialization through Stories of Disaster in Engineering Laboratories." *Social Studies of Science* 49 (6): 817–838. https://doi.org/10.1177/0306312719880266.