Technische Universität Dortmund

Fakultät für Erziehungswissenschaft, Psychologie und Soziologie

Transformational Governance of Academic Teaching

Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.)

vorgelegt von Sabine Lauer

geboren am 25.05.1980 in München

Erstgutachter: Prof. Dr. Uwe Wilkesmann

Zweitgutachterin: Prof. Dr. Liudvika Leišytė

Mai 2019

Content

Summary

This cumulative dissertation will conclude with a methodological reflection of a three-year research project on the topic of the 'transformational governance of academic teaching'. Given that methodological reflections of third-party funded projects have been mostly neglected, this concluding discussion will explain how textbook accounts on research designs and research methods have been applied in practice. The four publications originate from the research project entitled "TeachGov – Transformational Governance of Academic Teaching", which was funded by the Federal Ministry of Education and Research (BMBF; Bundesministerium für Bildung und Forschung) and ran between October 2013 and September 2016.

Based on the previous findings and the assumption of a continued cultural disregard for academic teaching, this project has investigated how universities can systematically upgrade the status of academic teaching beyond monetary incentives. The distinction between transformational and transactional governance was chosen for the analytical framework. In terms of teaching, transformational governance encompasses all organizational efforts that contribute to the establishment of a teaching culture in which individual teaching commitment and engagement is significantly valued, while frustrating factors in terms of teaching organization and infrastructure are minimized. In contrast, the new steering instruments that were introduced within the course of the New Public Management reforms can be subsumed under transactional governance.

The two leading research questions were as follows: (1) How (and why at all) do German universities try to raise the status of academic teaching in addition to monetary incentives? (2) How do different modes of teaching governance affect individual teaching behavior?

To answer the first research question, a qualitative approach was adopted to identify institutional arrangements of supposedly 'effective' mode of teaching governance in more depth. Therefore, in 2014, a multiple-case study at four universities that were among the 10 winners of the 'Competition for Teaching Excellence' was conducted. A total of 21 semi-structured expert interviews with various stakeholders (members of the rectorate, outstanding teaching practitioners, heads of center for teaching and learning,

and project managers) holding privileged knowledge in regards to the awarded teaching initiatives were conducted. The publications Schmid and Lauer (2016) and Lauer and Wilkesmann (2017) originate from this phase.

To answer the second research question—that is, if such institutional arrangements of supposedly 'effective' modes of teaching governance also affect individual teaching behavior—a quantitative approach was chosen to draw more representative conclusions regarding their effectiveness. Consequently, two online surveys were conducted. A follow-up survey was conducted in May 2015 targeting the professoriate at the four teaching-awarded universities, where another four not-awarded universities could be recruited to participate in the survey.

Additionally, a complete survey of all German university professors was able to be realized, which further allowed the analysis of potential influences of transformational governance modes on individual teaching behavior on a larger scale. The data collection took place between October 2016 and June 2017, where an overall response rate of 12.6% (n=2,663) was achieved. The findings reported in Wilkesmann and Lauer (2018) and Lauer and Wilkesmann (2019) were largely derived from this data.

A cursory overview of the four published articles follows.

In Schmid and Lauer (2016), the potency of individual actors within institutionalized settings at the four teaching-awarded universities was foregrounded. In delineating and theorizing the 'transformational' strategies and measures practiced by the award-winning universities, the interviews clearly indicated at the very beginning that single actors were instrumental in promoting teaching initiatives before they would eventually spread throughout the university. The empirical observations could be best conceptualized within the theoretical framework of 'institutional entrepreneurship' (DiMaggio, 1988) to determine how some deviant professors act as institutional teaching entrepreneurs, whereas their colleagues remain indifferent. By drawing on Emirbayer and Mische's (1998) multi-dimensional concept of human agency—which differentiates between the iterative (past), practical-evaluative (present), and projective (future) dimension—it could be shown in four exemplary vignettes that most of the outstanding teaching advocates in the sample had a professional background as

managers in private sector companies. This managerial habitus (iterative dimension) prompted them to not accept unsatisfactory or dysfunctional situations in their daily teaching (practical-evaluative), but rather take the initiative and actively push towards change (projective dimension). As a management consultancy for university rectorates, this article proposes a triple-s-strategy, to systematically scout/select, support and scale such individual efforts for the benefit of the organization university.

In contrast to the actor-centric perspective adopted in Schmid and Lauer (2016), Lauer and Wilkesmann (2017) highlight the governance perspective. In particular, the two modes of governance of academic teaching (transformational and transactional) were linked to the organizational learning discourse. In drawing on Argyris and Schön's (1978) distinction between single and double-loop learning, two transformational strategies of best-practice that have been used by these universities to achieve teaching excellence allowed for an interpretation as double-loop learning. In delineating two exemplary cases of double-loop learning concerning the university-wide implementation of new teaching formats as part of their institutional strategies to teaching excellence, the following research question was foregrounded: which kind of governance is required to manage such double-loop learning processes. In the two best-practice cases (one university of applied sciences and one research university), both forms of governance are required for the management of double-loop learning. In the case of a top-down instigation, transformational governance is especially required in terms of idealized influence and inspirational motivation. Intellectual stimulation becomes more important in the case of a more bottom-up trigger of organizational learning. Finally, in both cases, transactional governance was required for the university-wide implementation of new teaching routines.

With the help of a survey from 2016–2017 and 2009 from the previous research projects, Wilkesmann and Lauer (2018) examined if the influence of teaching motivation on the importance attached to methods of instructional design has changed with the implementation of New Public Management. Using self-determination theory and the concept of transformational and transactional governance, three hypotheses were tested via regression analyses. Whereas intrinsic motivation is the strongest predictor for the importance attached to methods of instructional design in both surveys, the influence of identified teaching motivation has only become statistically relevant

since 2016–2017. Albeit weak in strength, the transactional and transformational modes of governance also gained influence in 2016–2017. In contrast, a comparison of means reveals that feelings of guilt when neglecting one's teaching duties have considerably increased from 2009 to 2016–2017, while more autonomous forms (intrinsic, identified) of teaching motivation remain unchanged.

The relationship between the institutional environment and collegial exchange about teaching at German research universities was examined in Lauer and Wilkesmann (2019). Based on the nationwide survey from 2016–2017, a total of seven hypotheses were tested using regression analysis. These findings showed that professors significantly report a more vivid collegial exchange about teaching practices in the presence of a culture where colleagues hold teaching in relatively high esteem and when students actively provide constructive feedback. The university-wide governance of academic teaching also mattered. The perceived number of occasions where teaching is explicitly valued by the rectorate, the perception of how teaching is valued by the head of departments, and the perceived number of didactic support measures positively affected collegial dialogue about teaching. In contrast, no effect could be found for the general perception of how teaching is valued by the rectorate or for the number of perceived structural incentives.

Zusammenfassung

Thema dieser abschließenden Diskussion im Kumulus meiner Dissertationsschrift ist eine Methodenreflektion eines dreijährigen Forschungsprojektes zum Thema ‚transformationaler Governance akademischer Lehre'. Da methodologische Reflektionen von Drittmittelprojekten bis dato weitestgehend vernachlässigt worden sind, soll im Rahmen dieser abschließenden Diskussion nun dargelegt werden, wie Lehrbuchkonzepte zu Forschungsdesigns und Forschungsmethoden in der Praxis angewendet worden sind. Die vier Publikationen stammen aus dem Forschungsprojekt „TeachGov – Transformational Governance of Academic Teaching", das zwischen Oktober 2013 und September 2016 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert wurde.

Basierend auf früherer Forschung und der Annahme einer anhaltenden kulturellen Geringschätzung der akademischen Lehre wurde in diesem Projekt untersucht, wie Universitäten den Status der akademischen Lehre jenseits monetärer Incentivierung systematisch erhöhen können. Als analytischer Rahmen wurde die Unterscheidung zwischen transformationaler und transaktionaler Governance gewählt. In Bezug auf die Steuerung akademischer Lehre umfasst die transformationale Governance sämtliche organisationalen Anstrengungen, die zur Etablierung einer Lehrkultur beitragen. Das individuelle Lehrengagement wird geschätzt und gefördert, indem frustrierende Faktoren in Bezug auf Lehrorganisation und -infrastruktur minimiert werden. Im Gegensatz dazu können die neuen Steuerungsinstrumente, die im Rahmen der New Public Management Reformen eingeführt wurden, unter transaktionaler Governance subsumiert werden. Im Vordergrund dieses Projekts standen die folgenden zwei Forschungsfragen: (1) Wie (und warum überhaupt) versuchen deutsche Universitäten jenseits monetärer Anreize den Status der akademischen Lehre zu erhöhen? (2) Wie beeinflussen verschiedene Formen der Lehr-Governance das individuelle Lehrverhalten?

Um die erste Forschungsfrage zu beantworten, wurde ein qualitativer Ansatz gewählt. Auf diese Weise wurden institutionelle Arrangements von vermeintlich ‚effektiven' Formen der Lehr-Governance genauer untersucht. Im Jahr 2014 wurden hierzu Fallstudien an insgesamt vier Hochschulen durchgeführt, die im Rahmen des

Wettbewerbs ‚Exzellente Lehre' vom Stifterverband ausgezeichnet wurden. Die Hauptdatenquelle bildeten 21 semi-strukturierte Experteninterviews, die in unterschiedlichen Akteursgruppen (Mitglieder des Rektorats, herausragende Lehrpersönlichkeiten, Leiterinnen und Leitern von didaktischen Zentren und Qualitätsmanagementbeauftragte) durchgeführt wurden und deren Expertise sich jeweils über ihr privilegiertes Wissen bezüglich der preisgekrönten Lehrstrategien auszeichnete. Die Publikationen Schmid und Lauer (2016) und Lauer und Wilkesmann (2017) stammen aus dieser Projektphase.

Zur Beantwortung der zweiten Forschungsfrage wurde ein quantitativer Ansatz gewählt, um repräsentativere Rückschlüsse auf die Wirksamkeit transformationaler Wirkweisen der Lehr-Governance zu ziehen. Zu diesem Zweck wurden zwei Online-Umfragen durchgeführt: Eine im Mai 2015 durchgeführte Folgebefragung richtete sich an Professorinnen und Professoren an den vier lehrprämierten Universitäten, bei der vier weitere, für ihre Lehre nicht ausgezeichnete Universitäten für die Befragung gewonnen werden konnten. Darüber hinaus konnte schließlich eine Vollerhebung aller deutschen Universitätsprofessorinnen und -professoren nach Projektende realisiert werden, die es erlaubte, potentielle Einflüsse transformationaler Governance-Modi auf das individuelle Lehrverhalten anhand einer größeren Fallzahl zu analysieren. Die Datenerhebung fand zwischen Oktober 2016 und Juni 2017 statt, wobei eine Rücklaufquote von 12,6% (n=2,663) erzielt werden konnte. Die in Wilkesmann und Lauer (2018) sowie Lauer und Wilkesmann (2019) berichteten Ergebnisse stammen weitgehend aus dieser Befragung.

Bei Schmid und Lauer (2016) stand die Wirkmächtigkeit einzelner Akteure an den vier ausgezeichneten Universitäten im Vordergrund. Bei der Beschreibung und Theoretisierung der von den lehr-ausgezeichneten Universitäten implementierten ‚transformationalen' Strategien und Maßnahmen hat sich bei den Interviews von Anfang an herauskristallisiert, dass es hauptsächlich der Verdienst einzelner Akteure war, diese ausgezeichneten Lehrstrategien in Gang zu setzen, bevor diese sich dann schließlich über die gesamte Universität ausgebreitet haben. Diese empirischen Beobachtungen lassen sich am besten im theoretischen Rahmen von ‚Institutional Entrepreneurship' (DiMaggio, 1988) begreifen. So können Antworten darauf gefunden werden, warum einige ‚deviante' Professorinnen und Professoren als institutionelle Lehrunternehmer agieren, während andere Kollegen gleichgültig bleiben. Anhand von

Emirbayer und Misches (1998) relationalem Konzept der *human agency*, welches zwischen der *iterative* (Vergangenheit), der *practical-evaluative* (Gegenwart) und der *projective* (Zukunft) *dimension* unterscheidet, konnte in vier exemplarischen Vignetten gezeigt werden, dass die meisten dieser institutionellen Lehrunternehmer keine klassische akademische ‚Schornsteinkarriere' durchlaufen hatten, sondern vor ihrer Rückkehr an die Universität leitende Managementposten innehatten. Dieser manageriale Habitus (*iterative*) veranlasste sie, sich in ihrem täglichen Lehrumfeld nicht mit unbefriedigenden oder dysfunktionalen Situationen abzufinden (*practical-evaluative*), sondern Initiative zu ergreifen und maßgeblich an positiven Veränderungsprozessen mitzuwirken, welche den Stellenwert der Lehre an diesen Hochschulen deutlich erhöhte (*projective*). Als Conclusio für Universitätsrektorate schlägt dieser Artikel eine ‚triple-s-strategy' (*scouting*, *supporting*, *scaling*) vor, um solche individuellen Bemühungen systematisch zu identifizieren, zu unterstützen und bestenfalls hochschulweit zu implementieren.

Im Gegensatz zu der in Schmid und Lauer (2016) gewählten akteurszentrierten Perspektive wurde in Lauer und Wilkesmann (2017) die Governance-Perspektive betont. Die beiden Governance-Modi der akademischen Lehre (transformational und transaktional) wurden mit dem Diskurs des organisationalen Lernens verknüpft. Anhand der Unterscheidung von Argyris und Schön (1978) zwischen single-loop learning and double-loop learning konnten zwei der insgesamt vier ausgezeichneten Lehrstrategien als double-loop learning Prozesse identifiziert werden. Konkret beinhaltete diese double-loop learning Prozesse an beiden Hochschulen die hochschulweite Implementierung neuer Lehrformate. Folgende Forschungsfrage stand im Vordergrund: Welche Art von Governance ist erforderlich, um solche double-loop learning Prozesse zu managen?

Für beide Best-Practice-Hochschulen (eine Fachhochschule und eine Universität) konnte gezeigt werden, dass sowohl transformationale als auch transaktionale Governance erforderlich ist, um den organisationalen Lernprozess zu steuern. Am Beispiel der Fachhochschule, an der die Initiative stark top-down getrieben war, waren insbesondere folgende Aspekte der transformationalen Governance notwendig: *idealized influence* und *inspirational motivation*. Im Falle eines stärker bottom-up getriggerten Auslösers des organisationalen Lernprozesses, wie am Beispiel der

Universität gezeigt werden konnte, war der Aspekt der *intellectual stimulation* stärker ausgeprägt. Schließlich waren in beiden Fällen auch verstärkt transaktionale Modi der Lehr-Governance für die universitätsweite Umsetzung neuer Lehrformate erforderlich.

Wilkesmann und Lauer (2018) untersuchten anhand der Befragungen aus den Jahren 2016-2017 und 2009, ob sich der Einfluss der Lehrmotivation auf das Lehrengagement mit der Einführung von New Public Management (NPM) verändert hat. Unter Verwendung der Selbstbestimmungstheorie und des Konzepts der transformalen und transaktionalen Governance wurden mittels Regressionsanalysen drei Hypothesen getestet. Als zentrales Ergebnis lässt sich festhalten, dass die Wichtigkeit der Lehrmethodik nach wie vor im höchsten Grade selbstbestimmt motiviert ist. Dennoch lässt ein Vergleich der beiden Datensätze erkennen, dass nicht nur die identifizierte Lehrmotivation an Einfluss gewonnen hat, sondern auch Gefühle der Schuld und des Unwohlseins seitens der Professorenschaft verspürt werden, wenn die Lehrverpflichtungen vernachlässigt werden. Es lassen sich zudem erste positive – wenn auch schwache – Effekte einer transformationalen Governance auf die Wichtigkeit der Lehrmethodik empirisch nachweisen, nämlich organisationale Anstrengungen, eine ‚supportive teaching culture' (Feldman & Paulsen, 1999) zu etablieren. Wie erwartet sind selektive Anreize des NPM nach wie vor kein Anlass für Professorinnen und Professoren, der Lehrmethodik mehr Relevanz zuzuschreiben.

In Lauer und Wilkesmann (2019) wurden schließlich Einflussfaktoren auf den kollegialen Austausch über Lehre untersucht. Basierend auf der Befragung von 2016-2017 wurden mittels Regressionsanalyse insgesamt sieben Hypothesen getestet, die auf Feldman und Paulsens (1999) ‚supportive teaching culture' fußen. Die Ergebnisse zeigten, dass Professorinnen und Professoren signifikant häufiger über einen lebhafteren kollegialen Austausch über Lehrmethoden berichten, wenn das unmittelbare Lehrklima durch eine relativ hohe Wertschätzung von Lehre seitens der Kollegen wahrgenommen wird und sich Studierende durch konstruktives Feedback in der Lehre einbringen. Es konnten aber auch Einflüsse transformationaler Governance festgestellt werden: Sowohl die wahrgenommene Anzahl von Anlässen, bei denen Lehre explizit vom Rektorat wertgeschätzt wird, die Wahrnehmung, wie Lehre auf Fachbereichsebene wertgeschätzt wird und die wahrgenommene Anzahl von didaktischen Unterstützungsmaßnahmen, wirkten sich ebenfalls positiv auf den kollegialen Austausch über Lehrmethoden aus.

1 Introduction: A methodological reflection of a research project on 'Transformational Governance of Academic Teaching'

The main objective of this summarizing review for my cumulative dissertation on the topic "Transformational Governance of Academic Teaching" is to present a methodological reflection on a three-year research project[1] of the same name, which was funded by the Federal Ministry of Education and Research (BMBF; Bundesministerium für Bildung und Forschung) and ran between October 2013 and September 2016. This is the last part of a seven-year multimethod investigation into the 'managerial self-governance'[2] (de Boer, Enders, & Schimank, 2007) dimension of higher education institutions, which targets the governance of academic teaching at German universities where a total of three third-party funded projects have been realized.

Before moving on to the outline for the methodological reflection, I would like to give a brief overview of the theoretical and conceptual background to locate this research project within the overall research agenda.

The two preceding DFG projects[3] laid important groundwork for the successful acquisition of this project. These earlier projects had examined the influence of the new steering instruments (performance-related payment, performance-related budgets and management-by-objectives) that were introduced in the course of New Public Management (NPM) on academic teaching of German professors. The main results of these projects were that both individual teaching behavior and teaching motivation remained largely unaffected by these reforms (Wilkesmann & Schmid, 2011, 2012, 2014). To test their hypotheses, a teaching inventory was developed based on Cashin's

---

[1] (Project: 01PY13001, Funding period: 2013–2016)

[2] In essence, the governance perspective abandons the idea of a direct state control and focuses on "regulatory structures of a social unit and the interplay of these regulation structures" (Hüther & Krücken, 2018, p. 99). The 'managerial self-governance' is one governance dimension out of five (the others are: state regulation, external guidance, academic self-governance, and competition) of the governance equalizer which was developed with the intent to analyze how recent higher education policy changes towards New Public Management (NPM) affect different higher education governance regimes (de Boer, Enders, & Schimank, 2007, pp. 138–140).

[3] (Project: WI 2052/2-1, Funding period: 2008–2010; Project: WI 2052/2-2, Funding period: 2010-2012)

(1989) expanded definition of teaching, which covered the following aspects of everyday teaching: 'preparation and revision of content', 'methods of instructional design' and 'evaluation' (Wilkesmann & Schmid, 2012, p. 41). While assuming a rather cultural disregard for academic teaching, the theoretical underpinnings included principal-agent theory (Arrow, 1985; Eisenhardt, 1989a) and self-determination theory (Ryan & Deci, 2000a, 2000b). These authors concluded that teaching is mostly perceived as a highly self-determined activity, where the perception of the institutional environment hardly has any explanatory power. Additionally, a more student-focused approach to teaching (Trigwell, Prosser, & Waterhouse, 1999) and being female were further positive predictors for individual teaching behavior (Wilkesmann & Schmid, 2011, 2012).

Based on these findings, this last project "Transformational Governance of Academic Teaching", where the four publications contributing to this cumulative dissertation are largely based, has investigated how universities can systematically upgrade the status of academic teaching beyond monetary incentives. Consequently, Wilkesmann (2013) introduced the distinction between transactional and transformational governance, where the terminology was adapted from the full range leadership model (Bass & Avolio, 1993). In this context, these attributes are not understood as personal leadership styles but as structural and cultural features, hence the term governance is used (Wilkesmann, 2019, p. 112).

In essence, transactional governance mainly operates in the form of "explicit or implicit contractual relationships" (Bass & Avolio, 1993, p.116) by setting objectives and monitoring the respective outcomes. For the concrete case of academic teaching, prime examples of transactional governance can be seen in the selective incentives that were introduced during the NPM reforms, which were investigated in the two previous projects. In contrast, transformational governance covers the cultural dimension. It aims to create "a sense of purpose and a feeling of family" (Bass & Riggio, 2006, p. 3) within the organization and also establish "a rich set of norms which cover a wide range of behaviors, norms that will adapt to and change with external changes in the organizations' environment" (Bass & Avolio, 1993, p. 118).

In terms of teaching, transformational governance encompasses all of the organizational efforts that contribute to establishing a teaching culture in which individual teaching commitment and engagement is significantly valued, while frustrating factors in terms of teaching organization and infrastructure are minimized. A normative blueprint for this teaching culture is presented in Feldman and Paulsen's (1999) "supportive teaching culture", which combines both types of these governance modes.

The main objective of this project was to investigate the effectiveness of these transformational governance modes. Consequently, the following two central research questions were to be answered in a two-stage research design:

RQ1) How (and why at all) do German universities try to raise the status of academic teaching in addition to monetary incentives? *Which measures have been conceptualized and implemented? And, what are the experiences with their implementation?*

RQ2) How are these measures perceived by the professoriate? Are these measures in any form related to individual teaching behavior?

According to the research proposal, the initial research agenda had two main phases—a qualitative-explorative phase and a quantitative-explanatory phase:

1) To answer RQ1, our research proposal provided a qualitative approach that included a multiple-case study of four best-practice, teaching-awarded state-controlled universities. The major data collection took place at these universities between January and May 2014. A multiple-case study design was particularly suited to identify institutional arrangements of a supposedly 'effective' mode of teaching governance in more depth. Therefore, multiple data sources and perspectives were able to be included and triangulated. The publications Schmid and Lauer (2016) and Lauer and Wilkesmann (2017) originate from this phase. The findings from the German case were also contrasted with different higher education regimes, including the Netherlands, Hong Kong, and the UK. The data collection in these universities took place between June 2014 and 2014.

However, because the international comparison[4] was not used in the four articles contributing to this dissertation, it will not be addressed further here.

2) To answer RQ2, a follow-up survey was carried out at the respective four best-practice universities in May 2015, where the sample included another four universities that were not considered as best-practice institutions. In this second step, a survey design allowed us to investigate if the supposedly 'effective' modes of teaching governance that were identified in the first stage also affected individual teaching behavior. In other words, to draw more representative conclusions regarding their effectiveness. However, the findings obtained from this phase have not yet been published.

Another nationwide survey targeting all professors at German state-controlled research universities was conducted, which was not part of the initial research proposal. Hence, the data collection could only take place after the end of the official funding period, which took place between December 2016 and June 2017. In addition to an in-depth investigation of the influence of transformational governance on teaching-related behavior, we were also able to compare identical thematic blocks—such as those addressing transactional governance and teaching motivation—with a survey conducted with the same target audience in 2009. The findings reported in Wilkesmann and Lauer (2018) and Lauer and Wilkesmann (2019) were largely derived from this survey.

In the following, I will pass on to the methodology of the overall research process in review and I will shed more light on how the two research questions were methodologically and practically approached. Up to now, methodological reflections in the case of third-party funded projects have largely been neglected. Therefore, this summarizing discussion will close this research gap. As Hunter and Brewer (2015) conclude, design can have multiple meanings:

> On the one hand, [it can] be the outcome of prescriptive planning as in research proposals […]. On the other hand, design may be discerned in post hoc pattern recognition of what has been an unfolding, evolutionary, pragmatic adaption in the research process. In short, design can consciously occur before starting the research, or the process can occur as the research unfolds in an organic, opportunistic, and serendipitous fashion. (p.186)

---

[4] For a comparison of Hong Kong, the UK, and Germany see Wilkesmann (2016).

Using textbook accounts on research designs and methods, I will discuss how they were ultimately applied and adapted in the research process itself. Therefore, context and theory-related aspects and peculiarities of the field of study are largely omitted because both are well documented in the respective publications.

The rest of this summarizing review is structured as follows. First, each step of the qualitative phase will be methodologically reflected and contextualized with textbook accounts on the adapted research design and used methods. Second, the rationale of implementing a survey research as part of the second phase will be discussed in light of alternative research designs. Subsequently, the previously unpublished results from this phase will be presented. Third, a concrete example of the nationwide survey that has been realized after the project was already over will be given, and the general advantages and disadvantages of conducting online surveys will be discussed in regards to the target population of full professors. In the last section, I will conclude with an overall stocktaking of the research project in terms of methodology and key findings, where I will also point to dynamics that can be addressed in future research.

2   Level of involvement

Prior to moving to the main section, the following table clarifies my involvement throughout the research process.

Table 1 Division of labor in the research process

Preparatory work	
Problem formulation and research proposal	Uwe Wilkesmann and Christian J. Schmid
Pre-selection of the units of analysis	Uwe Wilkesmann and Christian J. Schmid
Phase 1: Multiple-case study	
Preparation and coordination of interviews	Sabine Lauer and Christian J. Schmid
Data collection	Uwe Wilkesmann, Christian J. Schmid, and Sabine Lauer
Data analysis	Sabine Lauer
Phase 2: Follow-up survey	
Designing the questionnaire	Sabine Lauer, Uwe Wilkesmann, and Alfredo Virgillito
Data analysis	Sabine Lauer
Phase 3: The nationwide survey	
Designing the questionnaire	Sabine Lauer, Uwe Wilkesmann, and Miriam Schmitt
Data analysis	Sabine Lauer

3   Phase 1: A multiple-case study of best-practice universities

As listed in Table 1, the practical implementation of the research project was carried out in a team. In addition to the conception of the research proposal, Uwe Wilkesmann and Christian J. Schmid were also significantly responsible for the data collection in the first phase. Although I was involved in the sense that we visited the sites together, my role in the interviews was clearly defined as being the statistician and novice (higher education) researcher who will mainly be responsible for the quantitative phase. On the one hand, this special role allowed me to fully concentrate on the discussions between the interview partners and the leading project researchers. On the other hand, I was ideally suited at that time to observe and reflect on the interviews in a relatively unbiased manner. However, as my knowledge and understanding of the project issues grew, I also began to ask more during subsequent interviews.

Before going into the practicalities of the first phase of the research project, I will briefly outline some of the basic ideas of the selected research strategy adopted during this phase.

3.1 A brief overview of case study research designs

In general, case studies constitute "a specific way of collecting, organizing, and analyzing data; in that sense, it represents an analysis process" (Patton, 2002, p. 447), where the term case study can simultaneously refer to "the process of inquiry (how we study the case) and the product of that inquiry (the write-up or representation of the case)" (Tobin, 2010, p. 771). However, definitions vary considerably across disciplines and fields of study (Schwandt & Gates, 2018), some of the most prominent follow:

> A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. (Yin, 2003, p. 13)

> Case study is not a methodological choice but a choice of what is to be studied. (Stake, 2005, p. 443)

Meanwhile, the literature has provided several typologies of case studies, where two main methodological positions can be distinguished (Yazan, 2015; Harrison, Birks,

Franklin, & Mills, 2017): a post-positivist orientation with its most prominent advocate Yin (2003) and a social constructivist/interpretivist orientation which is followed by Stake (1995, 2005, 2006).

Yin's (2003) typology describes three types of case studies: descriptive, exploratory (i.e., theory building or theory development), or explanatory (hypothesis and theory testing). An alternative typology is given by Stake (2005), who differentiates between intrinsic and instrumental case studies, and argues that an intrinsic case study "is undertaken, because […] one wants better understanding of a particular case" (p. 445) for its own values; whereas an instrumental case study seeks to provide "insight into an issue or to redraw a generalization" (p. 445). Consequently, approaches to case study designs can considerably vary in regards to their theoretical contribution and claims for generalization (Thomas, 2011; Ridder, 2017).

According to Yin (2003) "case studies are the preferred strategy when 'how' or 'why' questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context" (p. 1). What constitutes a case is usually not a trivial question. For example, a case can constitute a relatively bounded object or a process: it can be theoretical, empirical, or both (Ragin, 1992). At a minimum, a case can be defined as a phenomenon that is bounded in time and space. Additionally, the phenomenon on which a case study focuses may change during the research process, where "cases may be multiple in a given piece of research: What the case is may change both in the hands of the researcher (during the course of research and when the results are presented) and in the hand of the researcher's audiences" (Ragin 1992, p. 8). In addition, Yin (2003) suggests "as a general guide, your tentative definition of the unit of analysis (and therefore the case) is related to the way you have defined your initial research questions" (p. 23).

Case studies may include a single unit of analysis (single case study) or they may include multiple units of analysis (multiple-case study). A single case study aims to provide an in-depth understanding of the phenomenon under investigation, whereas a multiple-case study can provide "the researcher an even deeper understanding of processes and outcome of cases, the chance to test (not just develop) hypotheses, and a good picture of locally grounded causality" (Miles & Huberman, 1994, p. 26). In terms

of generalizability, Yin (2003) suggests that

> […] case studies, like experiments, are generalizable to theoretical propositions and not to populations or universes. In this sense, the case study, like the experiment, does not represent a 'sample' and in doing a case study, your goal will be to expand and generalize theories (analytic generalization) and not to enumerate frequencies (statistical generalization). (p. 10)

Consequently, when conducting multiple-case studies, his main argument is "*to consider multiple cases as one would consider multiple experiments*—that is, to follow a 'replication' logic" (Yin, 2003, p. 47). Accordingly, cases should be selected to predict either "similar results (*a literal replication*)" (p. 47) or "contrasting results but for anticipatable reasons (*a theoretical replication*)" (p. 47). In contrast, Stake (2006) takes the opposite view, where "both case studies and multicase studies are usually studies of particularization more than generalization […] the power of case study is its attention to a local situation, not in how it represents other cases in general" (p. 8).

In organizational studies, case studies are particularly well suited for the analysis of "the context and processes which illuminate the theoretical issues being studied. The phenomenon is not isolated from its context […] but is of interest precisely because the aim is to understand how behavior and/or processes are influenced by, and influence context" (Hartley, 2004, p. 323). However, since the case definition in organizational research is usually not orientated towards individuals but towards social processes, it may not be easy to define the analytical unit (and its boundaries). Therefore, the context has to be adequately described to understand the setting in which the case is examined.

In general, case study research is based on multiple data sources, most of which are qualitative and can be triangulated with each other. Meanwhile, interviews are the most important data collection strategy (Yin, 2003, p. 89). In addition to interviews, documents and observations are also commonly used in combination. Triangulation is often seen as a way to improve construct validity, where each source of evidence can be tested against another (Denzin, 1989; Flick, 2008). Accordingly, Pflüger, Pongratz, and Trinczek (2010) emphasize that an essential feature of case studies in organizational research is the systematic inclusion of different perspectives (*multiperspectivity*) because social processes are always the result of the interaction between several actors (or groups).

In principle, any data collection technique can be used, even if case studies are often misrepresented as qualitative methods (Eisenhardt, 1989b). Although Yin (2003) states that case studies "can include, and even be limited to, quantitative evidence" (p. 19), Robson and McCartan (2016) observe, that "it is relatively rare to see case studies where any quantitative component has anything other than a minor role" (p. 151). Consequently, the mixed-methods community (Creswell, 2007; Teddlie & Tashakkori, 2009) treats case study research as a distinct form of qualitative research, alongside ethnography, narrative investigation, grounded theory, and phenomenology.

Given that this research project is part of a larger research agenda into the governance of academic teaching, it may benefit from the previously developed theoretical propositions that were generated in the earlier studies. In this sense, we mostly followed the basic idea behind the case study approach by Yin (2003). Therefore, our main goal was to seek further theoretical explanations concerning the "how" and "why" questions of making teaching count more in German academia. However, in the further course of case study planning, data collection and data evaluation, we no longer explicitly orientated ourselves to the case study literature but instead consulted general handbooks for qualitative research. In alignment with Tight (2010), it can be concluded that

> Case study as a form of social research is not a particularly meaningful term. This is further evidenced when you examine the textbooks produced by advocates of case study. For, once you trip out the specific discussion of case study, which is typically restricted to not much more than an examination of specialist typology in the first chapter or two, what you are left with is a generic guide on how to do social research, usually with the emphasis on qualitative research. (p. 337)

Thus, by considering various sources of data, which primarily drew on in-depth interviews (supplemented by documents and participant observation) and included multiple perspectives (members of the rectorate, outstanding teaching practitioners, and higher education management professionals), our leading research question was as follows:

How (and why at all) do German universities try to raise the status of academic teaching in addition to monetary incentives?

The following section will provide insights into the research process; that is, how the multiple-case study was methodologically informed in practice, including sampling procedures, further preparatory work, data collection, data analysis, and a summary of the published results.

3.2 Selection of cases

In comparison to quantitative research, most sampling in qualitative research includes purposive sampling techniques, according to criteria that allow the research questions to be answered (Miles & Huberman, 1994; Patton, 2002; Yin, 2003). A comprehensive typology of different purposive sampling rationales is provided by Teddlie and Yu (2007), where the distinction is drawn between the following criteria: sampling to achieve representativeness or comparability, sampling special or unique cases, sequential sampling, and sampling using combinations of purposive techniques (as shown in Table 2).

Table 2 Purposive sampling techniques, adopted from Teddlie and Yu (2007, p. 81)

A. Sampling to achieve representativeness or comparability
1. Typical case sampling
2. Extreme or deviant case sampling (also known as outlier sampling)
3. Intensity sampling
4. Maximum variation sampling
5. Homogeneous sampling
6. Reputational case sampling
B. Sampling special or unique cases
7. Revelatory case sampling
8. Critical case sampling
9. Sampling politically important cases
10. Complete collection (also known as criterion sampling)
C. Sequential sampling
11. Theoretical sampling (also known as theory-based sampling)
12. Confirming and disconfirming cases
13. Opportunistic sampling (also known as emergent sampling)
14. Snowball sampling (also known as chain sampling)
D. Sampling using combinations of purposive techniques

The sampling strategy that was adopted in this multiple-case study can be best described as extreme case sampling; that is, cases where chosen that are the far ends of a particular dimension of interest. This purposive sampling strategy "involves selecting those cases that are the most outstanding successes or failures to some topic of interest" (Teddlie & Yu, 2007, p. 81). Thus, it is expected that such extreme successes or failures will provide particularly valuable information on the subject of interest.

At the time when the research proposal was in preparation, the first challenge to identify universities that already have implemented extraordinary teaching supportive frameworks. In contrast to research, no consistent metrics have yet been developed that allow a ranking of universities in regards to their teaching quality or teaching performance. Eventually, the search benefitted from the 'Competition for Teaching Excellence' (CTE, Wettbewerb Exzellente Lehre) (Brockerhoff, Stensaker, & Huisman, 2014). This competition was launched in 2009 by the Donors' Association for the Promotion of Science and Humanities in Germany (*Stifterverband*) and the Conference of Ministers of Education (*Kultusministerkonferenz*). This was the first competition where state-controlled universities could compete against each other in regards to whatever they defined as 'teaching excellence'[5]. Both, research universities and universities of applied sciences were eligible to participate. To enter the competition, the universities had to develop institutional teaching strategies that were based on their self-image and the previous achievements that would significantly increase their visibility and attractiveness as (excellent) teaching institutions, especially for undergraduate students. The concepts could focus on individual, particularly committed and high-performance departments/faculties of the respective university. However, the concept had to make the strategic importance of these subject-related activities clear for the development of the university as a whole, where only one concept could be submitted per university (Kultusministerkonferenz, 2008).

Out of 108 proposals submitted, 10 universities were awarded with funds of up to one million euro each to put their 'excellent' teaching strategies into practice. The winning universities included six research universities and four universities of applied sciences,

---

[5] Indeed, this teaching excellence initiative can be defined as a 'low fidelity' approach to teaching excellence that grants a high degree of definitional freedom, which Saunders (2009) sets in contrast to 'high fidelity' approaches that are more in alignment with "the prescription and conformity of the requirements of a specific policy or strategy" (as cited in Land & Gordon, 2005, p. 16).

where the both different university types were evaluated separately due to their organizational differences in regards to teaching and research.

For practical reasons, the number of cases was limited to four out of the 10 teaching-awarded universities, where a total of n = 2 research universities and n = 2 universities of applied sciences were purposively selected. Thus, in addition to the literal replication logic proposed by Yin (2003), the sampling strategy was additionally informed by the theoretical replication logic as contrasting results were anticipated in regards to university type.

As shown in Figure 1, a further selection criterion for the sample of research universities was that they should also exhibit a high degree of transactional governance. On the basis of a survey conducted in 2011, which addressed all rectorates of German universities, the degree of the respective university's transformational governance was computed as an average index of selected items relating to the rectorate's attitudes and assessments to various forms of indicator-based university management (Schmid & Wilkesmann, 2015; Wilkesmann, 2015).
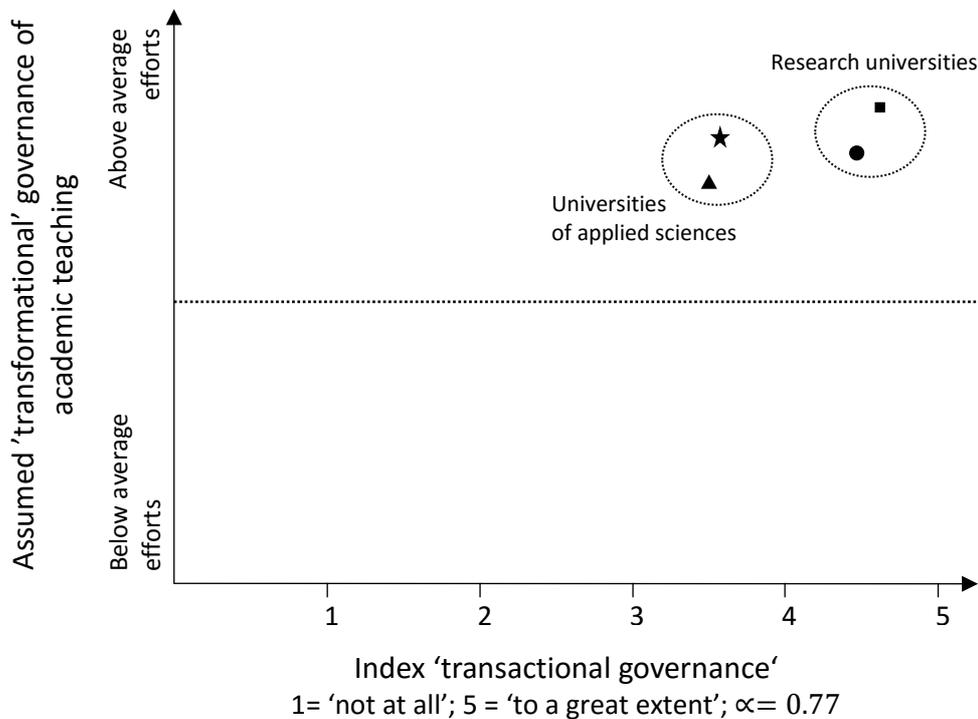
Figure 1 Sampling of extreme cases

3.3 Preparatory stage

In regards to the field access, the initial contact with the selected universities turned out to be rather unproblematic because three signed letters of intent to support the research project were already available when the research proposal was submitted. These letters of intent that were attached to the research proposal, and included both the assurance that a total of five interviews could be conducted and that an online survey among the professors would be send out via an email distribution list at a later stage of the project. Nonetheless, it is not self-evident that the access to the field is always as unproblematic as it was for the purpose of this study—at least for the first phase of the project. In general, the following six reactions can occur when organizations are approached for research purposes: *disinterest*, *fear* (being compared to other organizations or that organizational secrecies being disclosed), *tired of being subject to research*, *user-orientated* (weighing of the benefits), *recognition-seeking*, and *interest* (Pflüger, Pongratz, & Trinczek 2017, pp. 399–400). While it can be assumed that the selected universities had a general interest in our research project, the search for recognition could have been another motivational factor for the universities of applied sciences because they are not yet as over-researched as research universities (Wilkesmann, 2019).

Shortly after the start of the research project in October 2013, we had the opportunity to conduct an interview with the project manager of the CTE to obtain further background information about the award-winning universities and the competition itself. Given that this interview has not been published before and was indispensable for the success of the interviews, this methodological reflection is a fitting opportunity to share a few quotes.

In total, the interview lasted over two hours, where the project manager provided intimate insights of each university's prevailing teaching culture and pointed out the key-drivers of the respective teaching strategies. Therefore, we were already sensitized in advance to the fact that each university had their own 'story' and approach to teaching excellence ranging from a top-down, bottom-up to a combination of top-down and bottom-up strategy.

In Table 3, the four quotes describe the project manager's impression of two research universities and one university of applied sciences. Quotes (1) and (2) refer to a research university with a strong top-down approach to teaching excellence, quote (3) refers to a research university with a combined top-down/bottom-up approach to teaching excellence, and quote (4) refers to a university of applied sciences with a bottom-up approach to teaching excellence:

Table 3 Exemplary quotes on the teaching-awarded universities

(1)	When you read the application, you notice that this university has already written many applications. And they attach a great deal of importance to the overall governance of teaching, so this proposal reads. But I have to be honest, I didn't find the proposal really sympathetic, it's not necessarily such a proposal where your heart rises […] because unlike the other proposals, it's so far away from the student. So, one does not read between the lines that the students are important to them. You get the feeling they want to manage it well, but ...
(2)	Don't get me wrong, but in comparison to many other universities, at this university I had less the impression that the concept really reaches teaching faculty and students. During the inspection there are many representatives of administrative departments. Prior to that, there was the ceremony for the teaching awards, okay there were people around […] but shortly afterwards, there were no students and there were no lecturers around anymore, only one member of administrative staff, that had something lifeless, as if the university only consisted of the organization.
(3)	Out of the four universities, this university perhaps followed the most bold strategy that impressed the most with its SWOT analysis […] and they also distributed it very broadly over many shoulders, but unlike as in University X, this is not only distributed to administration staff, but on the professors as well who are always responsible for the subprojects […] and they involved a lot of people. And at no university were so many university members present at the day of the inspection, students, and teachers as well as university staff […] although both research universities are very similar universities, the strategic concept of University Y is way more alive, not least because of its storybook project management.
(4)	This university has a full house on the day of teaching, a very high percentage of professors are participating. On this day, the teacher takes part and they have filled the largest lecture hall of the university. I could get a picture of it there. The rector is there for the whole day, so there is a high commitment of the whole university but which was originally created and bundled by these two key drivers.

Thus, in the interview with the project manager of the CTE, we were disillusioned even before the official interviews began that there was apparently no exclusively top-down organizational solution to systematically upgrade the status of teaching at universities. The same holds true for a consolidated definition of teaching excellence: Within the framework of the CET, the award-winning universities were further obliged to participate in a quality circle. The aim of this quality circle was twofold. On the one hand, the universities should exchange information on the different institutional teaching strategies; while on the other hand, this exchange should also bear fruit in the creation of a joint 'Charter for Good Teaching' (*Charta gute Lehre*) (Jorzik, 2013). However, rather than a common definition of what constitutes good or excellent teaching, the result of this Charter for Good Teaching was rather disappointing from the point of view of the promoters because it contained, for the most part, a collection of partly contradictory best-practice solutions.

3.4 Planning the expert interviews

In the next step, the interviews were coordinated via email, where we contacted the respective rectorates with a brief reminder of our research project and the request to recommend further interview partners. Although the project manager of the Stifterverband had already pointed out the main drivers of the award-winning strategies, we were still looking for further interviewees who played a decisive role in the award-winning strategies or were otherwise highly committed teaching advocates worth talking to. Therefore, expert interviews seemed to be a suitable method of data collection to investigate how the teaching-awarded universities are trying to systematically upgrade the status of their teaching. Accordingly, each within-case sample should primarily include *experts* from various status groups holding privileged knowledge in regards to the awarded teaching initiatives; that is, actors who can be made responsible for the design, implementation or control of a problem solution and/or who has privileged access to information about groups of individuals or decision-making processes (Meuser & Nagel, 2005, p. 73). Thus, the underlying definition and status of experts is characterized by their social position and their professional role[6] (e.g., as members of a particular organization or group of professionals) which provides

---

[6] A less restrictive definition would be to regard everyone as experts in their own life stories (cf. Gläser & Laudel, 2010).

them privileged access to separate fields of knowledge.

Bogner and Menz (2009) distinguish between three forms of expert knowledge: *technical knowledge*, (2) *process knowledge*, and (3) *interpretative knowledge*, as follows.

(1) Technical knowledge "contains information about operations and events governed by rules, application routines that are specific to the field."

(2) Process knowledge "relates to the inspection of and acquisition of information about sequences of action, interaction routines, organizational constellations, and past or current events."

(3) Interpretative knowledge includes "the expert's subjective orientations, rules, points of view and interpretations, which suggest a picture of expert knowledge as a heterogeneous conglomeration." (p. 52)

However, expert knowledge can only be identified as interpretative knowledge when the data is collected and the data analyses are carried out. In this sense, this kind of expert knowledge is always an abstraction and systematization achievement of the researcher, an "analytical construction" (Bogner & Menz, 2009, p. 53).

Accordingly, Bogner and Menz (2009) distinguish the following types of expert interviews: an exploratory expert interview, a systematizing expert interview, and a theory-generating expert interview. If the researcher has little knowledge about the field that they want to investigate, then an exploratory expert interview is best suited for "sounding out the subject under investigation" (Bogner & Menz, 2009, p. 46). Systematizing expert interviews focus on "knowledge of action and experience which has been derived from practice, is reflexively accessible, and can be spontaneously communicated" (pp. 46–47). This type of expert interview is based on a definition of the expert as a person holding special knowledge about social facts (Gläser & Laudel, 2010, p. 12), where the expert's "function is rather that of informants who provide information about the real objects being investigated" (Bogner & Menz, 2009, p. 47). In contrast, the theory-generating expert interview abandons the notion of the expert being a provider of specific knowledge, instead "its goal is the communicative opening up and analytic reconstruction of the subjective dimension of expert knowledge" (p. 48).

According to this typology, only the theory-generating expert interview is capable to grasp "the social relevance of expert knowledge" (p. 53) then.

In this sense, who counts as an expert is not based on personal characteristics and special skills and abilities, but is a construct from the researcher; it "remains a relational concept [...] and has to be answered in relation to the concrete field of operation in which the expert acts, and with reference to the investigative spectrum of the empirical study being carried out" (Bogner & Menz, 2009, p. 54). An approximate methodological definition follows:

> An expert has technical, process and interpretative knowledge that refers to a specific field of action, by virtue of the fact that the expert acts in a relevant way (for example, in a particular organizational field or the expert's own professional area). In this respect, expert knowledge consists not only of systematized, reflexively accessible knowledge relating to a specialized subject or field, but also has to a considerable extent the character of practical or action knowledge, which incorporates a range of quite disparate maxims for action, individual rules of decision, collective orientations and patterns of social interpretation. An expert's knowledge, his or her orientations and so on, also (and this is the decisive) point to the fact that she or he may become hegemonic in terms of practice in his or her field of action (for example, in a certain organizational-functional context). In other words, the possibility exists that the expert may be able to get his or her orientations enforced (at least in parts). As the expert's knowledge has an effect on practice, it structures the conditions of action of other actors in the expert's field in a relevant way. (Bogner & Menz, 2009, p. 55)

In accordance with the case study literature, the authors equally stress that further data sources and methods must be used to answer the question of the social effectiveness of the expert knowledge. Therefore, the search for traces of expert knowledge was designed to be multi-perspective, and a total of the following three status groups were examined:

- Organizational perspective: vice-rectors/presidents for teaching and learning (n=4);
- Individual perspective: Exceptional and/or award-winning teaching staff, i.e. full professors (n=9) and n=1 senior councilor (*akademischer Oberrat*);
- Administrative (technical and didactic support) perspective: higher education professionals (n=7 heads of center for teaching and learning, e-learning, project, and quality managers).

With regard to Bogner and Menz's (2009) typology of expert interviews, the interviews conducted in this project can be defined as a mixture of systematizing expert interviews and theory-generating expert interviews. In agreement with the literature, the interviews were relatively open and informed by a topic guide to provide each interviewee with enough room to unfold their own outlooks and reflections. Consequently, the comparability of expert statements was methodically ensured in the topic guide and empirically ensured by the shared institutional environment of the interview partners (Meuser & Nagel, 2009).

3.5 Academics conducting expert interviews with academics

Two epistemological problems will be briefly discussed here, which must generally be reflected in higher education research. A peculiarity of higher education research is that it is conducted by academics and thus by members of higher education institutions. Consequently, academics not only conduct research about their own organization but also about their own field (Wilkesmann, 2019, pp. 39–50). According to Wilkesmann (2019), the following two interconnected epistemological problems can occur: the "self-objectivation problem" (Rindermann, 2000) and the "overestimation problem" (Wilkesmann, 2017). The first problem may result from the fact that different status groups may have different interests and self-assertion powers in carrying out their research. At worst, this can lead to evidence-based practical implications based on hidden interest-driven policies. This problem can only be overcome "through critical reflection, through recognition that one's research results may well be shaped by one's position in the power structure and by the ideological context within which one carries out social scientific activities" (Sjoberg, Williams, Vaughan, & Sjoberg, 1991, p. 36). The second problem occurs when the researched population perceives itself as alleged experts in higher education research. This can often lead to premature, inadmissible generalizations that exclusively relate to one's own discipline. However, within the framework of this research project, the first problem could be ruled out because it was primarily concerned with the advancement of scientific knowledge and there were no hidden interests on the side of the research team. At a later stage in our research project, we were asked by the Stifterverband whether we would like to take over the evaluation of the CTE. It was precisely this conflict between supposedly value-neutral research and

an evaluation mandate that finally led us to reject their request. The overestimation problem was also not an issue in the interviews because the interview partners were genuinely interested in the overall research topic. Had this problem occurred, however, one could have reacted directly in the immediate interview situation.

Against this background, it goes without saying that, due to the procedural character and the non-explicitness of essential parts of the expert knowledge, it is unlikely that the expert interview follows a knowledge extraction process in the sense of a questionnaire. Rather, experts will report on cases of decision making in reference to the maxims that guided their action, where "these are the data necessary for constructing the supra-individual, field-specific patterns of expert knowledge" (Meuser & Nagel, 2009, p. 31). Hence, to get access to this exclusive knowledge, the expert interview requires the interviewer not only act upon methodological guideline but to become an intervening variable in the interview situation him or herself. These so called "interaction effects should be treated not as variables that distort the situations, but as elements, which are constitutive of the process of data collection" (Bogner & Menz, 2009, p. 45). Thus, the statements made in relation to the interviewer are likely to be influenced by the perceptions and attributions towards the interviewer which can be differentiated along the following three dimensions:

(1) Attributions in terms of professional competence and disciplinary field,
(2) Assumptions about the normative judgements on the field and the role of the interviewee and if the experts assume shared or diverging normative orientations,
(3) Assumptions about the interviewer's potential to influence and the attributed consequences to the interview (Bogner, Littig, & Menz, 2014, p. 51).

In total, Bogner and Menz (2009, pp. 55–69) differentiate six ways that the status of the interviewer may be ascribed by the interviewee in the interview situation: (1) the interviewer as co-expert, (2) the interviewer as expert from a different knowledge culture, (3) the interviewer as lay-person, (4) the interviewer as authority, (5) the interviewer as accomplice, and (6) the interviewer as potential critic.

The literature on expert interviews usually prefers the perception of the interviewer as co-expert, which is propagated as an ideal type for the interview situation 'at eye level' (Pfadenhauer, 2009). In this case, the interviewer is regarded as an equal partner and colleague with whom the expert exchanges knowledge and information. Due to a shared knowledge base, this interview constellation is supposed to be best suited in explorative or systematizing expert interviews, when the primary purpose is to gather information and elucidate facts. Because teaching is one of the core tasks of university professors, it is obvious that an expert interview on this subject matter between two equally ranked academics (as in the case of the interviews within the status group of organizational decision/policy makers and proven experts in teaching) can be best described as a conversation between two experts, where the interviewer is either perceived as a co-expert or as an expert from different knowledge culture. In theory-generating expert interviews, Bogner and Menz (2009) argue that this interview constellation may become problematic, as "the implicit normative and practical premises of expert opinion will be presupposed as a shared basis of the conversation between expert and co-expert, and it will difficult to gain access to them for the purposes of analysis" (p. 60). However, to find answers to the 'how' and 'why' these universities succeeded to make teaching count more at their institutions, it were precisely these normative and practical premises ('why')—besides gathering information about the technical and procedural aspects of the respective teaching strategies ('how')—that were the focus of our research interest.

3.6 Data analysis

Based on the topology of the expert interviews, two approaches for the analysis of expert interviews are proposed in the literature (Bogner, Littig, & Menz, 2014): qualitative content analysis (Gläser & Laudel, 2010), which is used for systematizing interviews; and a grounded theory approach (Meuser & Nagel, 2005), which is used for theory-generating expert interviews. According to Gläser and Laudel (2010), the goal of qualitative content analysis of expert interviews is to reconstruct social facts. In this context, the aim is to find out the truth about certain states or processes. Thus, the focus is on attempting to uncover causal dependencies through a systematic analysis and comparison of the information provided by the experts. Therefore, they propose a modified version of Mayring's (2014, 2015) structured content analysis, with a deductive category assignment that follows a five steps procedure (Gläser & Laudel,

2010, pp. 197–260). As in all content analytical procedures, the centerpiece is an ex ante determined category system that is derived from theoretical considerations and prior research findings, which will then be applied to the data. In fact, the categories must prove themselves on the material but they must be open enough to allow for new categories to arise. Gläser and Laudel's (2010) modification of Mayring's (2014, 2015) approach now consists in keeping the categories as open as possible to maintain the tension between theory and data.

In contrast, the analysis of theory-generating expert interviews does not focus on gathering information or reveal the objective truth about certain states or processes but focuses instead on attempting to tap into the interpretative knowledge of the experts; that is, to identify the guiding principles, rules and values that significantly determine the thinking and interpretation of the experts (Bogner, Littig, & Menz, 2014, p. 76). Accordingly, the goal of the data analysis strategy proposed by Meuser and Nagel (2005), which is based on grounded theory reasonings, is to work out the "supra-individual commonness [own translation]" (Meuser & Nagel, 2005, p. 80) of the expert interpretations by comparing the available expert interviews; that is, to reconstruct jointly shared, typical knowledge stocks, relevance structures and patterns of interpretation. Their suggested data analysis approach follows four steps: coding, thematic comparison, sociological conceptualization, theoretical generalization. Consequently, the focus is on thematic units, which account for the importance of "the institutional-organizational context within which the expert's position is embedded and which provides the actor with guiding principles" (Meuser & Nagel, 2009, p. 35). In contrast to the deductive approach, which tries to subsume the data into preconceived concepts or theories, this grounded theory approach develops plausible and theoretically sophisticated readings of the expert statements derived from the data. This should grant the researcher with the greatest possible openness of diverse possibilities of interpretation. Nevertheless, Bogner, Littig and Menz (2014) warn against the illusion of theoretical impartiality on the part of the researcher, especially when analyzing theory-generating expert interviews. The early phase of coding and category building is already likely to be influenced by theoretical concepts. However, these concepts and resulting patterns of interpretations should be understood as sensitizing concepts (Meuser & Nagel, 2005, p. 90), which provide researchers "directions along which to

look", but "which lack precise reference and have no bench marks which allow a clean-cut identification of a specific instance and of its content" (Blumer, 1954, p. 7).

Consequently, neither of these two data analysis strategies has been applied in a textbook manner because our approach was informed by both. Although the starting points for analyzing the data were predetermined coding schemes, the coding schemes themselves were kept parsimonious and straight-forward. Consequently, we mostly followed a 'directed content analysis' (Hsieh & Shannon, 2005) where those parts of the interview transcripts were highlighted, which seemed to be related to the predetermined coding scheme but which instead consisted of two overall sensitizing concepts: institutional entrepreneurship and governance of academic teaching. In other words, transformational (idealized influence, individualized consideration, inspirational motivation, and intellectual stimulation) and transactional (selective incentives, monitoring) modes. No additional themes emerged inductively while analyzing the data. However, the interview with the Stifterverband had sensitized us for an alternative conceptual framework; that is, in addition to the organizational perspective, one must also take a strong actor-centric perspective. Consequently, without this initial interview, the concept of 'institutional teaching entrepreneurship' may have only emerged in the course of the data analysis. Because the first interviews took place at a university whose award-winning teaching strategy was purely a bottom-up achievement, the idea that some kind of 'institutional teaching entrepreneurship' is already at work manifested as sensitizing concept at the very beginning of the interviews. This significantly led to a sharper focus on this concept in the subsequent interviews.

3.7 Findings

In Schmid and Lauer (2016), we used Emirbayer and Mische's (1998) relational definition of human agency as coding scheme where each university of the interviews within the status group of proven experts in teaching could be analyzed based on the sensitizing concept of 'institutional entrepreneurship' (DiMaggio, 1988). Here, the focus of analysis was mainly to elaborate the interpretative knowledge of how and why these actors became so instrumental in the awarded initiatives; that is, how and why their orientations became "hegemonic in terms of practice in his or her field of action" (Bogner & Menz, 2009, p. 55). Accordingly, the data analysis consisted of the

systematic exploration of their interpretative knowledge in regards to the following three dimensions of human agency: the iterational dimension (*past*), the practical-evaluative dimension (*present*), and the projective dimension (*future*) (Emirbayer & Mische, 1998). In regards to the iterational dimension, our key finding is that these actors have all spent a considerable amount of time outside academia, mostly in senior management positions. While it is mandatory to spend at least five years outside academia to apply for a professorship at a university of applied sciences, most of the academic careers at research universities are pursued within academia. This shared iterational dimension strongly draws from a managerial and entrepreneurial habitus and this informed their practical-evaluative dimension in this sense that they encountered the prevailing teaching culture at their universities as problematic. This led them to act upon their incorporated maxims and routines, which are different to those of the average professor whose incorporated maxims and preferences were exclusively shaped in academia.

In our conclusion and management advice, we proposed a 'triple s-strategy,' which included the following three elements: scouting, supporting and scaling. Scouting refers to the appointment procedure; that is, the systematic recruitment of suitable candidates who bring with them the essential characteristics of an institutional teaching entrepreneur. Therefore, the focus is on the iterational element. Supporting refers to the practical-evaluative dimension of their human agency. Meanwhile, scaling applies to the projective dimension. On the one hand, rectorates must ensure that their efforts become more visible and effective; on the other hand, they also have to grant expert status to them (e.g., as delegate of the rectorate or even as a full member of the rectorate) to secure their future agency.

In the second step, in Lauer and Wilkesmann (2017) we adopted an organizational learning perspective[7] where the distinction between transformational and transactional

---

[7] However, it has to be noted, that analyzing the data from an organizational learning perspective was more of an opportunity than an initial 'theoretical' data analysis strategy as the second author of this article was asked if he would like to contribute to a special issue on organizational learning in higher education. To get a comprehensive overview of various approaches of organizational learning in higher education institutions, the systematization provided in Dee and Leisyte (2016) is highly recommended.

governance was used to analyze how "double-loop learning" (Argyris & Schön, 1978) in teaching and learning can be managed. Within the sample of the four teaching-awarded universities, two institutional teaching strategies could be identified as a serious commitment to engage into a double-loop learning process, which involved "the modification of an organization's underlying norms, policies and objectives" (Argyris & Schön, 1978, p. 3) in regards to their teaching mission. In concrete terms, both strategic concepts that were awarded within the CTE comprised the university-wide implementation of new teaching formats to improve their overall teaching performance. In the case of the university of applied sciences, every curriculum had to implement mandatory teaching formats that were based on problem-based learning. In contrast, the strategic concept of the research university was strongly orientated towards the university-wide spread of digital teaching and learning formats. The interviews with the respective vice-rectors for teaching and learning showed how the interplay of transformational and transactional governance interacted in the concrete implementation process of these new teaching formats. Consequently, the different components of transformational governance could mostly reconstructed by referring to the interpretative knowledge that the vice-presidents for teaching and learning attached to the university's and their overall motivation to push the teaching agenda forward at their institutions. For triangulation purposes, the interviews from the institutional entrepreneurs were also part of the data analysis.

To gain insight into the process knowledge which addressed the management component of these double-loop learning processes, the data was then analyzed within the category of transactional governance. The data analysis once again substantiated the crucial role of the respective institutional teaching entrepreneurs that were identified in Schmid and Lauer (2016) because they were either the key-drivers of the awarded concepts, as in the case of the research university, or in the case of the university of applied sciences were otherwise instrumental in pushing the teaching reforms throughout the university.

In accordance with Ragin (1992), the phenomenon on which this multiple-case study initially focused allowed for multiple interpretations in the course of the data analysis. Out of the initial research question—'How can a university upgrade the status of teaching beyond monetary incentives?'—an actor-centric interpretation of the data

emerged where the case could be termed 'institutional teaching entrepreneurship at German universities.' Meanwhile, the organizational perspective allowed for an interpretation of some of the data as 'the governance of organizational learning in teaching and learning at German universities.'

In terms of plausibility and trustworthiness of the findings, especially if the strong emphasis on the professional background within the sample of the institutional teaching entrepreneurs is over-exaggerated, we did a member check in a subsequent conversation with one of the acclaimed institutional teaching entrepreneurs, who assured us that our theoretical interpretations were completely accurate. The two universities whose institutional strategies were framed as double-loop learning processes were eventually awarded as the first universities to receive the Genius Loci award from the Stifterverband, a national teaching award that was introduced in 2017 (Stifterverband, 2019). Finally, participating in didactic workshops and other events held by the respective teaching-awarded universities were also informative to help us to understand and validate the contextualities even more comprehensively than the formally scheduled expert interviews alone could provide (Schmid, 2016).

Our publication strategy included one journal article and one book chapter; however, only part of the data was eventually utilized for data analysis. In addition to the 'institutional teaching entrepreneurs', as portrayed in Schmid and Lauer (2016), we found further empirical evidence for similar types of teaching advocates who have been crucial in pushing the teaching agenda forward at their institutions—either in their immediate sphere of influence or on a larger scale. What these actors have in common with the institutional teaching entrepreneurs is that most of them also share an equivalent managerial habitus that is derived from outside academia. In addition, their extraordinary teaching commitment resulted out of problematic situations that they faced in their daily teaching. Consequently, some of these actors did not shy away from deviating from the existent guidelines or from formal communication channels.

4 Phase 2: The follow-up survey

The leading research question for the second phase of the research project addressed the effectiveness of transformational and transactional governance modes that were witnessed at the teaching-awarded universities to upgrade the status of academic teaching. In particular, our investigation focused on how these organizational endeavors are perceived by the professoriate. In a second step, we asked if they also affect individual teaching behavior.

Initially, the data analysis strategy intended to link both the qualitative and the quantitative phase in terms of a joint interpretation of both findings. In this sense, within the overall multimethod design of this research project, the two stages could have been framed as "exploratory sequential design" (Creswell, Plano Clark, Gutmann, & Hanson, 2003, p. 228). This is in the tradition of mixed-methods designs[8]:

> A mixed methods study involves the collection or analysis of both quantitative and/or qualitative data in a single study in which the data are collected concurrently or sequentially, are given priority, and involve the integration of the data at one or more stages in the process of research. (p. 212)

However, this definition does not apply to several single studies within a sustainable research program in which researchers can combine methods at different stages of the research. According to Morse (2003), the exploratory sequential design is characterized by an "inductive theoretical drive"; that is, the results from the qualitative inquiry are given priority. Therefore, this design is best suited where an in-depth exploration of the phenomenon under inquiry is required to add substance to a theoretical framework, to identify further important variables which have been previously disregarded, or to inform the development of measurement instruments (Creswell & Plano Clark, 2007, p. 75). Accordingly, the main objective of the multiple-case study was to extend the existing theoretical framework but at the same time was also on the lookout to identify further variables that might affect individual teaching behavior. In any case, the qualitative phase took precedence over the quantitative phase, although strictly speaking the order did not matter because each phase pursued its own research questions.

---

[8] For a detailed discussion on the methodology of mixed methods see e.g. Tashakkori and Teddlie (2003), Creswell and Plano Clark (2007), Kelle (2008) or Burzan (2016).

However, due to low response rates and a lack of sound operationalized dependent variables, the quantitative results have never been linked to the results derived from the multiple-case study in a peer-reviewed journal publication. Nevertheless, selected results of this quantitative phase should at least be presented within the framework of this methodological reflection.

4.1 Preparing the questionnaire

Given that the main goal of the qualitative phase was not the construction of a measurement instrument, document analysis was mainly used to inform the questionnaire design. Consequently, the homepages of the respective teaching-awarded universities were examined with regard to their supporting teaching-learning infrastructure. In particular, the following dimensions were distinguished:

- The existence of didactic support structures (various course offering in didactic training, special and/or mandatory training for newly appointed professors);
- The existence of selective incentives (e.g., teaching grants, reduction of teaching load, teaching sabbaticals);
- The existence of institutionalized platforms for collegial exchange about teaching (e.g., professional learning groups, teaching blogs, day of teaching).

Additionally, the survey included questions about the general appreciation of teaching by the rectorate, by the department, and by colleagues. It also asked for their perceptions of how teaching achievements are appreciated at various occasions, such as the day for teaching or by special mentions in print- and online media.

Items that were inspired from the findings in Schmid and Lauer (2016) addressed the reputation of particularly teaching-active colleagues and asked if the professors had occupied a former management position prior to their appointment as a professor.

How the results from student teaching evaluations are handled on the departmental level was also inspired from the empirical evidence derived from the interviews:

For instance, at one teaching-awarded research university, which is well-known for its fairly neo-liberal managerialism, a departmental ranking of individual teaching

evaluations was introduced and the rectorate utilized these results as an assessment of their performance. Those chairs that performed below average were 'symbolically' punished with a minimal shortening of their basic funds (cf. Schmid & Lauer, 2016). Consequently, an internal 'quasi' competition for teaching 'reputation' was established where "nobody wants to take a place down below [bottom of ranking]" (research university, professor). In such an environment, people are more likely to observe, compare and mimic each other's achievements, or even try to outclass colleagues by any means necessary (which also holds for teaching in this case).

Furthermore, we found empirical evidence at these two research universities that the deans used the student evaluation results as a constructive conflict management tool to challenge the professors to reflect on their teaching performance and find solutions for personal improvement. In this case, the respective dean confronts their colleagues with conspicuities regarding their teaching performance (e.g. poor student ratings, high dropout rates or simply student complaints) or their general 'marching to a different drummer' attitude towards their teaching duties. They then carefully ask for explanations and (in some cases) make some suggestions to change the deficiencies. However, a face-to-face conversation can be problematic when there are severe grounds, as one male professor from a research university succinctly puts it:

> There is so much ego involved, and that doesn't tolerate a lot of severe personal critique. When you criticise a method, this is still far away from the person, but when it comes to one's own teaching, it is always person-related, [...] and that is perceived as a personal attack. (research university, dean)

In addition, another dean speaks from experience when he says that it is generally difficult to face colleagues with a wagging finger "because as a dean you are not equipped with disciplinary power and once the term of office is over, you are equal ranked colleagues again". One of these universities deploys a 'special trick' in such crisis talks: the respective dean first gets the students' feedback and they then mediate between both parties, "that always works—the customer is king" (research university, dean).

On the individual level, the items measuring potential dependent variables were related to teaching motivation and teaching commitment, with discipline, gender, and tenure as control variables.

At the end, due to project member turnover and delays in designing the questionnaire, we were under time pressure to finish the survey. In some ways, this resulted in a lack of methodological rigor in regards to the operationalization of the dependent variable. In particular, teaching commitment was only measured with four items, as will be shown later. In particular, the retrospective should have used more reliable measurements, such as the previous developed inventories for teaching behavior (Wilkesmann & Schmid, 2011, 2012) and teaching style (Wilkesmann & Lauer, 2015).

4.2 Sample description

The follow-up survey was conducted in May 2015 and targeted all professors at the awarded universities. Additionally, another four universities which were not awarded for their teaching excellence were included to the sample. Because a pre-post comparison of the teaching-awarded universities in the sense of a panel design was not possible, this contrasting case selection was made to compensate for this deficiency. The invitations to participate in the survey were sent out by email by the respective rectorates. One disadvantage of this approach was that we did not know how many professors would eventually receive the invitation to the survey. Although the rectorates of the teaching-awarded universities agreed in the letter of intent to support the survey study at a later date, some universities agreed to forward the link to the survey only after repeated requests had been made. Therefore, it is doubtful if the basic population (n = 2,296) had actually been reached by the end of the survey. As shown in Table 4, a total of 321 professors participated in the survey, which leads to an overall response rate of 16.6%, where the response rate at the teaching-awarded universities was with 14.3%, which is slightly lower than at the not-awarded universities 19.0%. Unfortunately, the sample was unbalanced because only one not awarded research university could be recruited.

Table 4 Sample description I

University	Type	TA	University size	Response rate
1	UAS	yes	large	12,5%
2	UAS	yes	small	25,0%
4	RU	yes	large	11,6%
5	RU	yes	large	7,9%
3	UAS	no	small	20,2%
6	UAS	no	medium	30,3%
7	UAS	no	medium	10,7%
8	RU	no	medium	14,8%
				16.6%

UAS = university of applied sciences, RU = research university, TA = teaching-awarded university,
large = "more than 300 professors", medium = "150-300 professors", small = "less than 150 professors"

Further sample characteristics are shown in Table 5, where it has to be noted that the sample was not intended to be representative for the population of professors at German research universities and universities of applied sciences. Instead, it should have been representative for the purposive sample of these eight universities in regards to gender, payment scheme, and disciplines. However, this comparison would have required a special, fee-based analysis from the Federal Statistical Office (*Destatis*) because the freely available statistical metrics are only on the aggregate level. This special analysis would also have made little sense because it cannot be assumed that all professors received the invitation to the survey.

Table 5 Sample description II

Variable	% within sample	n
Male	72.2	228
Female	27.8	88
W-salary	66.7	214
C-salary	33.3	107
Research university	41.4	133
University of applied sciences	58.6	188
Soft disciplines		
Linguistics and cultural studies	13.4	43
Law, economics, and social sciences	20.3	65
Hard disciplines		
Mathematics and natural sciences	24.7	79
Engineering	29.7	95
Others	11.9	38

Nevertheless, it is still worthwhile to carry out selected statistical analyses. First, the awarded and not-awarded universities will be compared in regards to:

1. The perception of the overall appreciation of quality teaching on the university (president and vice president for teaching and learning) and departmental level (faculty dean, vice dean for teaching and learning, and colleagues);
2. The perception of how quality teaching is defined.

Second, a comparison will be made in regards to:

1. The perception of the university-wide visibility of quality teaching (teaching awards, teaching days, and mentions in print and online media);
2. The perception of the institutional teaching support (existence of didactic group training, individual training, peer coaching, peer communities, online support, and technical support);
3. The perception of how teaching quality is monitored (obligatory didactic training for newly appointed professors, teaching evaluations);
4. The perception of the structural incentives to promote quality teaching (teaching grants, sabbatical for teaching, reduction of teaching load).

4.3 Comparison of teaching-awarded and not awarded universities

As can be seen in Table 6, the perceived appreciation of 'quality teaching' by the rectorate is the highest at the teaching-awarded universities of applied sciences (M = 4.35, SD = 1.00), whereas it is the lowest at not-awarded universities of applied sciences (M = 3.6, SD = 1.24). In contrast, no significant differences can be found for the research universities. Furthermore, the remaining comparisons in regards to the appreciation of quality teaching by the department and the colleagues do not reveal any significant differences. In terms of the perceived institutional definition of 'quality teaching,' both types of universities show significant differences in favor to the teaching-awarded universities, where it is perceived highest again at the universities of applied sciences (M = 3.87, SD = 1.17), whereas it lower in the case of teaching-awarded research universities (M = 3.11, SD = 1.07).

Table 6 Average perception of the appreciation and institutional definition of quality teaching

	Research universities		Universities of applied sciences	
Appreciation of 'quality teaching' (1 = "strongly disagree", 5 = "strongly agree")	Not-awarded university (n = 37)	Teaching-awarded University (n = 94)	Not-awarded university (n = 121)	Teaching-awarded University (n = 67)
	M (SD)	M (SD)	M (SD)	M (SD)
Rectorate (president, vice president for teaching and learning)	3.70 (1.00)	3.84 (1.00)	3.58 (1.24)	4.35 (1.00)**
Department (faculty dean, vice dean for teaching and learning)	3.44 (1.34)	3.77 (1.01)	3.81 (1.19)	4.08 (1.18)
Colleagues	3.44 (1.25)	3.54 (0.97)	3.62 (0.99)	3.75 (1.07)
Institutional definition of 'quality teaching' (1 = "not at all 5 = "to a high degree")	2.40 (1.06)	3.11 (1.07) **	2.43 (1.11)	3.87 (1.17)**

** sig < 0.01; * sig < 0.05

In terms of visibility of 'good teaching practice' on the university level (Table 7), separate Chi-square tests for independence by university type show that significance differences can only be found for universities of applied sciences. Here, professors at teaching-awarded universities significantly report with 97.0% more often the existence of a teaching award ($\chi^2(1) = 13.79$, p = 0.00) than professors at not awarded universities.

The same holds true with 76.1% for the existence annual teaching day ($\chi^2(1) = 110.47$, $p = 0.00$) and with 74.2 % consents for special mentions in print- and online media ($\chi^2(1) = 16.84$, $p = 0.00$).

Table 7 Perception of the visibility of 'quality teaching' on the university level (% of yes answers)

	Research universities		Universities of applied sciences	
	Not-awarded university (n = 37)	Teaching-awarded University (n = 94)	Not-awarded university (n = 121)	Teaching-awarded University (n = 67)
Visibility of 'good teaching practice' on the university level				
Teaching award	89.2%	94.8%	76.0%	97.0%**
Teaching day	52.8%	52.6%	3.3%	76.1%**
Print and online media	64.9%	58.3%	42.9%	74.2%**

** sig < 0.01; * sig < 0.05

Regarding the perception of existent institutional support for teaching (Table 8), the data shows that teaching-awarded universities seem to always be significantly better equipped than not-awarded universities regardless of university type. When comparing the average number of didactic support offers, the perceived number of institutional support is with M = 3.90 (SD = 1.52) significantly higher at teaching-awarded universities than at not awarded universities (M = 2.35, SD = 1.41).

Table 8 Perception of the institutional support for teaching (% of yes answers)

	Research universities		Universities of applied sciences	
	Not-awarded university (n = 37)	Teaching-awarded university (n = 94)	Not-awarded university (n = 121)	Teaching-awarded university (n = 67)
Institutional support for teaching				
Group training	63.9%	87.5%**	81.0%	94.0%*
Individual training	25.7%	78.9%**	40.8%	83.3%**
Peer coaching	16.7%	28.1%	25.6%	77.6%**
Peer communities	22.2%	45.8%*	22.3%	76.1%**
Online support	25.7%	59.1%**	33.0%	62.5%**

** sig < 0.01; * sig < 0.05

In terms of monitoring teaching quality (Table 9), one way of signaling to new professors at the very beginning of their term of office that the university places a high emphasis on teaching is to oblige them to undergo obligatory didactic training. Indeed,

this is only the case at the teaching-awarded universities where 90.9% ($\chi^2(1) = 9.99$, p = 0.00) of the professors at the universities of applied sciences and still 24.0% ($\chi^2(1) = 10.45$, p = 0.00) of the professors at the research universities affirm its existence.

How student evaluations are handled on the departmental level differ significantly for the subsample of research universities, where 52.2% (($\chi^2(1) = 18.02$, p = 0.00) of the professors at the teaching-awarded universities report that this is an established practice. In contrast, only 13.5% of the professors at not awarded universities affirm this practice. No significant differences can be found in regards to the publication on student evaluation results within the department.

Table 9 Perception of how teaching quality is monitored (% of yes answers)

	Research universities		Universities of applied sciences	
	Not-awarded university (n = 37)	Teaching-awarded university (n = 94)	Not-awarded university (n = 121)	Teaching-awarded university (n = 67)
Monitoring of teaching quality				
Obligatory training for newly appointed professors	0%	24%**	70.8%	90.9%**
Teaching evaluation addressed by dean	13.5%	54.2%**	40.5%	40.3%
Teaching evaluation published	45.9%	40.6%	23.3%	31.3%

** sig < 0.01; * sig < 0.05

As can be seen in Table 10, the most frequent reported structural incentives are competitive teaching grants where significant differences can be found for both university types. At research universities, 83.3% professors of teaching-awarded universities significantly more often affirm the existence of teaching grants at their institution (83.3%), whereas the percentage of professors at not-awarded research universities is only 44.7%. At the universities of applied sciences, 65.7% of the professors from teaching-awarded universities of applied sciences report on their existence, where it is only 50.0% at not-awarded universities of applied sciences.

At 36.5% and 20.9%, the existence of a sabbatical for teaching is also significantly more often reported by professors at awarded research universities and universities of applied sciences. The less frequently reported structural incentive is a reduction of teaching load.

Table 10 Perception of the existence of structural incentives (% of yes answers)

	Research universities		Universities of applied sciences	
	Not-awarded university (n = 37)	Teaching-awarded university (n = 94)	Not-awarded university (n = 121)	Teaching-awarded university (n = 67)
Structural incentives to promote 'quality teaching'				
Teaching grants	44.4%	83.3%**	50.0%	65.7%*
Sabbatical for teaching	8.3%	36.5%**	8.3%	20.9%*
Reduction of teaching load	5.6%	7.3%	10.0%	13.6%

** sig < 0.01; * sig < 0.05

Especially in the qualitative phase, we have found which structural incentives work well in teaching and which do not. At first glance, the biggest incentive seems to be the sabbatical for teaching. However, similar to a teaching award, this is not a selective incentive in the strict sense because the motto here is that the winner takes it all (Wilkesmann & Schmid, 2012). The expert interviews with the rectorates also revealed that this structural incentive is not at all an incentive for professors who are already very committed to teaching because they like to teach and do not want to be released from teaching for one semester. Accordingly, teaching funds are regarded as the most important incentive because they are set at a much lower threshold and several professors can benefit from this incentive per semester. This is also an attempt to initiate an inner-university competition in teaching, which varied in strength across the four teaching-awarded universities in the sample.

In light of these cultural and structural differences, the following preliminary conclusion can be drawn: According to the perception of the surveyed professors, the awarded universities seem to have created a more privileged institutional environment than not-awarded universities. However, does a more 'supportive teaching culture' (Feldman & Paulsen, 1999) also have an effect on individual teaching behavior? This assumption will be investigated in the following two regression analyses.

4.4 Regression analyses

The teaching-related items in this survey asked the professors how much importance they attribute on a five-point Likert scale (1 = "not important at all"; 5 = "very important"), to the following aspects regarding their commitment to teaching:

- "In my teaching, I include a new didactic element every semester."
- "In my teaching, I'm constantly driven by testing out new things (e.g. online or IT-assisted tools)."
- "In my teaching, I'm always on the lookout for new impulses from my students."
- "In my teaching, I make sure to improve the social interactions between my students."

A principal component analysis with varimax rotation (KMO = .701, explained variance of 78.7%) was carried out and resulted in a one factor solution (Cronbach's Alpha = .77).

To measure the influence of a supportive teaching culture, the dummy variable teaching-awarded university (1 = yes, 0 = no) was used as a proxy. Teaching motivation was assessed with items from Fernet, Senécal, Guay, Marsh, and Dowson's (2008) work tasks motivation scale for teachers (WTMST). All of the items were measured on a five-point Likert scale (Wilkesmann & Schmid, 2012, 2014). A confirmatory factor analysis (cf. Appendix) based on the ULS estimator shows with GFI = 0.973, AFGI = 0.958; SRMR = 0.073 a still satisfactory model fit (Hu & Bentler, 1999). The respective Cronbach's alphas range from 0.61 to 0.71, which is also still in an acceptable range considering the short-scale character of the inventory.

In addition, control variables included discipline (1 = hard disciplines, 0 = rest), gender (1 = male, 0 = female), payment scheme (1 = W-salary, 0 = C-salary), and management experience outside academia (1 = yes, 0 = no).

Table 11 Regression analyses with dependent variable 'commitment to teaching'

	Research universities			Universities of applied sciences		
	B	*SE B*	*β*	*B*	*SE B*	*β*
Supportive teaching culture						
Teaching-awarded university (1 = yes, 0 = no)	0.01	0.17	.00	0.27	0.12	**.15***
Teaching Motivation						
Intrinsic motivation	0.39	0.12	**.37****	0.12	0.11	.09
Identified motivation	0.12	0.13	.10	0.24	0.11	**.19***
Introjected motivation	-0.03	0.07	-.04	0.06	0.06	.08
External motivation	-0.20	0.16	-.14	-0.11	0.12	**-.16***
Control variables						
Hard disciplines (1 = yes, 0 = no)	-0.02	0.16	-.01	-0.29	0.12	**-.17***
Gender (1 = male 0 = female).	-0.36	0.16	**-.19****	-0.43	0.13	**-.23****
Payment Scheme (1 = W-Salary, 0 = C-Salary).	0.25	0.15	.13	0.30	0.12	**.17***
Former management position outside academia (1 = yes, 0 = no)	0.01	0.16	.01	0.24	0.11	**.14***
n			129			179
adj. R²			0.23			0.25

** sig < 0.01; * sig < 0.05

As shown in Table 11, the model fit for the regression model for the sample of research universities totals 21.0% and 25.0% for the sample of universities of applied sciences. The results from the two separate regressions models reveal that teaching commitment is only significantly higher at teaching-awarded universities than at universities of applied sciences ($β$ = .15, p = .02). No difference can be found for the research universities in the sample. In regards to teaching motivation, the strongest positive predictor is intrinsic motivation ($β$ = .37, p = .00) at the research universities, where this type of teaching motivation remains insignificant at the universities of applied sciences ($β$ = .09, p = .28). However, identified motivation is only positively associated with

teaching commitment at universities of applied sciences (β = .19, p = .02) where no significant association can be found for research universities (β = .10, p = .38). In addition, external motivation is only negatively associated with teaching commitment at universities of applied sciences (β = -.17, p = .02). In regards to the control variables, professors from the hard disciplines significantly exhibit a lesser degree of teaching commitment (β = -.17, p = .01) at universities of applied sciences. Being male significantly decreases teaching commitment at both university types, whereas the payment scheme (β = .17, p = .01) and prior management experience outside academia (β = .14, p = .04) only have a positive effect at universities of applied sciences.

However, at this point a detailed discussion of the results is omitted because the results should be interpreted with caution due to the small sample size and the unbalanced distribution of the teaching-awarded and not awarded universities within the subsample of research universities. Nevertheless, at least for the subsample of universities of applied sciences, it can be stated that a more pronounced supportive teaching culture has small significant positive effects on the importance attached to individual teaching behavior.

5   Phase 3: The (unplanned) nationwide survey

In the last step, our research project should also generate added value for an evidence-based management of academic teaching. Therefore, a diagnostic tool was implemented where universities could get an assessment on their prevailing teaching culture. Because this was once again a welcome opportunity to conduct another nationwide survey, we designed a more detailed questionnaire to also do justice to our own research purposes. Unfortunately, this survey had to exclude professors at universities of applied sciences because a comprehensive database of their contact details is no longer available.

For the technical realization of the survey and the diagnostic tool, financial resources from the research project were still at hand to commission a company that specialized in employee surveys. The diagnostic tool's main function was that universities can get a personal report consisting of how the appreciation of teaching by various status groups (i.e., rectors/presidents, vice-rectors/presidents for teaching and learning, faculty deans, deans of studies and colleagues) is averagely perceived by the respective professoriate. Additionally, the perceived culture of appreciation in regards to teaching was assessed; that is, occasions where teaching is explicitly valued by the rectorate, such as at the annual day for teaching, special mentions in the university's print and online media, in personal conversations or announcements, acknowledgements and feedback, and guidelines and mission statements in the university development plan or quality teaching manual. In addition to the perceptions of the prevailing status of academic teaching, selected items in regards to the immediate teaching environment and infrastructure were also part of the report, including the student's characteristics, teaching infrastructure, and teaching climate. Finally, the personalized analysis also included selected characteristics of the professoriate, particularly occasions where they mostly reflect about teaching and their teaching motivation. Consequently, a university-specific evaluation was only possible if the respondent indicated their university affiliation.

5.1 Preparation of the survey

In comparison to the follow-survey conducted in 2015, this survey intended to be a complete survey of all full professors currently employed at German research universities. To recruit the participants, our chair purchased online access to the

'Deutsche Hochschullehrerverzeichnis' (Deutscher Hochschulverband, 2016). This database contains the contact details of all full professors who are currently employed at public and private research universities. To get access to the contact details, empty query statements could be made alphabetically. We were able to exclude retired, extraordinary professors and professors at private universities from the search. Accordingly, the collection of the email addresses followed alphabetically because there was no function to filter the respective professors by institution. This meant that in practice that two student assistants and myself ended up copying and pasting every single email into an Excel spreadsheet while keeping each other informed about which letter of the alphabet we were currently working on. This process was very time consuming, not least due to a sometimes very slow database where queries collapsed quite often.

In total, 22,405 email addresses were collected and the collection of the email addresses took about three months. In October 2016, the first email invitations to participate at the survey were sent out. A total of 21,089 emails could be successfully delivered until February 2017. Participation at the survey was possible between October 2016 and June 2017. As an incentive to participate, the professors were offered the opportunity to subscribe to a mailing list to gain access to the university-specific results generated by the diagnostic tool at a later date. These university-specific results were only available if at least five professors from the respective university participated. To increase the response rate, one further reminder was sent.

The questionnaire asked (similar to the previous survey) how academic teaching is appreciated, supported, and incentivized on the university and departmental level. Thus, these aspects address the transformational and transactional modes of teaching governance. We also re-assessed the teaching motivation and utilized the dimension "importance of teaching methods" from Wilkesmann and Schmid's (2011, 2012) teaching inventory. In alignment with the literature, which suggests that the maximal completion time of online surveys should not exceed 15 minutes (Engel & Schmid, 2019), we kept the survey short with an average completion time of 12 minutes. In comparison to the surveys conducted in 2009 and 2011, we passed on time-consuming questions, such as estimating percentages of the current workload for teaching, research, and administration, and estimating the number of exams and the number of students.

5.2 Response and nonresponse rates

Between October 2016 and June 2017, the link to the survey was visited 4,554 times and 2,663 questionnaires were fully completed. This leads to a response rate of 12.6% and a dropout rate of 9.0%. The number of professors who indicated their university affiliation totals 2,287 (85.9%). Although we have guaranteed the anonymity of the answers in our cover letter (i.e., the individual answers will not to be traced back to the person in our analyses), in the end it was up to each participant in the survey whether they wanted to feed the diagnostic tool with their participation or not. For the later statistical analyses, we then only included those respondents who also stated their university affiliation.

In general, two types of nonresponse can be differentiated: unit-nonresponse occurs, where the sampled unit does not respond to the request to the survey; and item-nonresponse, where particular items remain unanswered (Wilkesmann, 2019). While item-nonresponse depends to a remarkable extent on the survey questions and the overall questionnaire design, unit-nonresponse reflects a more general behavioral trend. As Menold and Züll (2011) state, denials represent a significant proportion of systematic failures in surveys. For refusals in face-to-face surveys, they could identify the following reasons, among others: general refusal (e.g. no interest or no time), negative attitude against surveys, and because of data protection (Menold & Züll, 2011, p. 102).

Ramm (2014), for example, notes that the willingness of German students to participate in a nationwide annual student survey has been declining since the 1980s, from an initial participation rate of 40.0% to less than 20.0% in 2012–2013. However, the problem of declining response rates is not only a specific problem in higher education research (Wilkesmann, 2019, pp. 77–81). Therefore, in light of these overall trends, a response rate of 12.6%, as was the case for this survey, is still in an acceptable range (Engel & Schmidt, 2019).

Even if the sample, except for minor deviations, represents the basic population quite well (Wilkesmann & Lauer, 2018; Lauer & Wilkesmann, 2019), one suspicion cannot be ruled out—our study might only include those professors who already hold teaching in a relatively high esteem. As can be seen in Table 12, this suspicion is further

substantiated when examining the single items of the teaching motivation inventory used in the survey, where professors were asked for the reasons why they teach, where the mean scores of the respective items related to intrinsic and identified teaching motivation range between M = 3.97 (SD = 1.02) and M = 4.41 (SD = 0.76).

Table 12 Teaching motivation (Wilkesmann & Lauer, 2018)

Items	Mean (SD)
Intrinsic teaching motivation (1 = not at all, 5=to a large extent)	
... because I derive much pleasure for my teaching.	4.25 (0.79)
... because I find teaching interesting.	4.13 (0.86)
... because I lose myself in teaching.	2.80 (1.14)
Identified teaching motivation (1 = not at all, 5 = to a large extent)	
.. because, for me, the task of teaching is of personal importance.	4.41 (0.76)
... because I see my teaching as a significant contribution to my students' overall academic progress.	4.25 (0.84)
... because the task of teaching provides the chance to realize an aspect of my academic profession that is of personal meaning to me.	3.97 (1.02)
Introjected teaching motivation (1 = not at all, 5 = to a large extent)	
... because my aspiration is also to be successful at teaching, otherwise I would feel like a loser.	3.48 (1.15)
... because I have a bad conscience if I have neglected my teaching duties.	3.84 (1.15)
.. because I would feel uncomfortable if I have neglected my teaching duties.	3.93 (1.01)
… because my self-concept as a professor is also determined by quality teaching.	4.21 (0.91)
External teaching motivation (1 = not at all, 5 = to a large extent)	
.. because my university/employment contract demands that I teach.	1.52 (0.89)
.. because I get paid for it.	1.37 (0.81)
Amotivation (1=not at all, 5=to a large extent)	
I don't know why, because the work conditions provided for academic teaching are unbearable.	1.61 (0.95)
I don't know why, sometimes I don't see the actual purpose of teaching.	1.68 (1.0)
I don't care much for teaching because I don't know what it effects.	1.36 (0.73)

For the concrete case of this survey it can be suspected then that professors who exhibit a lesser degree of dedication and commitment to teaching, and thus interest in the overall topic of this research project, might be underrepresented. Thus, the refusal to participate is suspected to be motivated by a general disinterest. However, we cannot verify this assumption afterwards. In comparison to the surveys conducted in 2009 and 2011, where the investigators of the last two research projects received many complaints via emails (Wilkesmann, 2019), the complaint emails in this survey were rather rare. Only a few professors sent emails where they decided not to participate due to reasons of data protection. Some examples of negative attitudes against surveys follow:

> Ich werde ganz bewußt Ihren Fragebogen nicht beantworten. Diese Art von Quantifizierung führt zu keiner Verbesserung und ist meist nur die Basis für weitere Quantifizierungsanliegen. Qualität kann man aber nicht zwingend in Zahlen messen, da neben den materiellen Werten auch häufig ideelle Werte gefragt sind. Diese gehen aber in derartigen Befragungen total unter. Im letzten Dialog des Phil. Brecht mit Koll. Lesch (ZDF am 27.11.2016) wurde dies besser herausgearbeitet als ich es auszudrücken vermag. Das Zählen und Bewerten von geistigen Leistungen (Ranking, Publikationen, Evaluationen...) ist eine Idee der Ökonomen und wird in stetig zunehmendem Maße angewandt. Es führt zu keiner neuen Qualität, zu keiner Innovation und zu keinem Fortschritt. Dieser Entwicklung schließe ich mich nicht an.
> (Email retrieved on November 29th, 2016)

Or

> Die Beantwortung Ihres Fragebogens habe ich abgebrochen, weil mir die Fragen teilweise (a) zu „kompetenzorientiert", d. h. inhaltsleer und fantasielos erschienen, (b) weil ich der Zusage der Anonymität nicht traue, (c) weil ich inflationäre empirische Befragungen leid bin und nicht erkennen kann, welche Erkenntnisse dadurch gewonnen werden, (d) weil sich Hochschulpolitiker und ideologische Kollegen der Ergebnisse solcher Studien gerne bedienen, um nichts ändern zu müssen und (e) weil die gewissenhafte Beantwortung aller Fragen weitaus mehr als nur zehn Minuten Zeit kostet. Im Übrigen wäre ich dankbar, wenn mir keine „Erinnerungen" ins Haus geschickt würden. Keine Antwort ist gerade im Zeitalter digitaler Diarrhöe auch eine Antwort.
> (Email retrieved on February 14th, 2017)

However, these thoughts were rarely expressed in the feedback emails. Some professors have even taken this survey as an opportunity to tell us about their own tales of suffering and woe, and how little teaching is valued their university, for example:

Ich bin ein leidenschaftlicher "Lehrer" (natürlich auch Forscher) an der Uni X und ärgere mich generell über die allerorten verbreitete Geringschätzung der akademischen Lehre seit vielen, vielen Jahren. Ich bin angetreten, Forschung UND LEHRE zu machen, nicht nur Forschung und notgedrungene Minimallehre. Meine Studierende kenne ich alle mit Namen und weiß oft auch so manches anderes von ihnen. Dies ist mir wichtig, denn ein System wie die Universität kann kein soziales sein, wenn alles nur laufende Nummern sind. Ich schätze die persönliche Begegnung mit den Menschen - den Studierenden wie den Kollegen - sehr hoch ein und ärgere mich jedes Mal, wenn ich Kollegen kennenlerne, die aus meiner Sicht so ganz ungeeignet als Vermittler sind bzw. generell die Lehre vernachlässigen und die Studierenden wie lästige Fliegen abfertigen. Ich kenne leider sehr, sehr viele Kollegen, die GENAU SO auftreten.
(Email retrieved on October 12th, 2016)

Ihr Forschungsprojekt ist ausgesprochen verdienstvoll. Da Sie mich unter meiner Adresse an der UNIVERSITÄT X angeschrieben haben, habe ich die Umfrage so ausgefüllt, als wäre ich noch Professor an der UNIVERSITÄT X. Dort war ich 12 Jahre lang tätig und bin das Gefühl nie losgeworden, mir gerade für meine Studierenden sämtliche Beine ausgerissen und dafür von der Hochschulleitung im Gegenzug - salopp gesprochen - nur auf die Fresse bekommen zu haben. Dementsprechend vernichtend sind in Bezug auf die UNIVERSITÄT X meine Antworten auf die Fragen in Ihrem Fragebogen ausgefallen, den ich soeben ausgefüllt habe. Ich sage es ganz deutlich: Wer als Hochschullehrer Wert darauf legt, die jungen Menschen, die ihm anvertraut sind, angemessen zu betreuen, ist an der UNIVERSITÄT X völlig fehl am Platz. Allen Sonntagsreden zum Trotz interessiert sich das Präsidium dort null komma null für die Lehre und für die Studierenden. An der UNIVERSITÄT X zählen nur Drittmittel und die schöne Außenfassade. Alles andere ist der Hochschulleitung dort egal. Tatsächlich bin ich seit dem 1. Oktober 2015 Professor an der UNIVERSITÄT Y, und wenn ich Ihren Fragebogen in dieser Eigenschaft ausgefüllt hätte, wären meine Antworten deutlich positiver ausgefallen. Das Arbeitsumfeld ist hier in Y für Professoren, die an der Lehre interessiert sind, um ein Vielfaches angenehmer.
(Email retrieved on November 11th, 2016)

Nonetheless, for future research, it would be conceivable, for example, to get to know in advance if there is any interest among the academic community to participate in the survey. This could be realized with an advance survey announcement that grants the option to be removed from the mailing list. Consequently, a link to a short optional questionnaire could be included that asks for the reasons for non-participation according to the reasons worked out by Menold and Züll (2011). On this occasion, other basic characteristics of the non-respondents could be also assessed. In the case of a teaching-related survey, this could include a short assessment of the respondent's teaching motivation or general preference for teaching, field of study, age, and gender. If this data would had been available in this survey, comparisons could have been made

between respondents and non-respondents to determine if our suspicion that the sample is biased towards more teaching-committed professors holds true. In practice, however, such short questionnaires that particularly address potential non-respondents are still rather rare (Weinhardt & Liebig, 2015).

5.3 Data analysis and results

In both Wilkesmann and Lauer (2018), and Lauer and Wilkesmann (2019), standard OLS regression analyses were carried out to test our hypotheses in regards to the importance of instructional design and perceived degree of collegial exchange about teaching. Although testing our hypotheses with a multilevel modeling approach (Snijders & Bosker, 1999) would have been an alternative data analysis strategy because over 80.0% of the respondents' indicated their university affiliation, single-level analyses were regarded as appropriate because of the low intra class correlation coefficients (ICC) and unequal group sizes of respondents per university (ranging from one to 61). Given the ICC = 1.56% for the dependent variable in Lauer and Wilkesmann (2019) and an ICC = 2.32% in Wilkesmann and Lauer (2018), these results clearly indicate that the largest variance of the respective dependent variables is caused by individual determinants, where the institutional environment hardly plays a role. In regards to the previous findings by Wilkesmann and Schmid (2011, 2012), the main results of these studies could be largely replicated: Teaching is and remains a highly self-determined activity, where Wilkesmann and Lauer (2018) could show that the new steering instruments that were introduced in the course of the NPM reforms still have not the desired effect on individual teaching behavior. A comparison of means shows that the importance attached to methods of instructional design has not significantly changed between 2009 (M = 3.45, SD = 0.79) and 2016–2017 (M = 3.41, SD = 0.81). In addition, the step-wise regression reveals that elements of a supportive teaching culture—such as the perceived didactic support, the extent of collegial exchange about teaching and constructive feedback from the students—gained significance but only add 3.0% more of explained variance to the base model, which explains 16.0% and includes teaching motivation and the control variables gender, discipline, and age. Further findings that could be replicated are the strong influence of discipline and gender: Placing emphasis on methods of instructional design is rather a matter for female professors and is considered more important in the soft disciplines.

As far as individual teaching motivation is concerned, probably the most exciting result originates from the comparison of the teaching motivation between 2009 and 2016–2017, where the introjected motivation to teach has significantly increased. Teaching has become more important for a professor's self-image and their sense of guilt has clearly increased if teaching duties are neglected. Still, at this point, it can only be suspected that this is due to the recent developments in higher education policy, which makes it plausible that quality teaching is now more on the agenda at universities than 10 years ago. Interestingly, the data shows no difference in regards to the introjected teaching motivation between older and younger professors, although these recent developments that promote teaching would suggest the rise of a new socialization effect especially among younger professors.

In Lauer and Wilkesmann (2019), the factors that affect collegial exchange about teaching were examined. Again, given the recent developments in higher education policy it is somewhat surprising that the professors' self-reported perception of their institutional environment with regard to an active collegial exchange about teaching (M = 2.39, SD = 1.12) has not changed over the last decade (M = 2.38, SD = 1.09) because the comparison of the data from 2009 and 2016–2017 confirmed. Nonetheless, the regression analysis revealed that professors report a more active collegial exchange about teaching when they are exposed to a departmental peer culture where teaching is held in high esteem and when students provide constructive feedback. Additionally, a supportive teaching culture is also positively related to more active collegial exchange about teaching. The peer culture and direct student interaction remain the most important predictors. This is also corroborated by the descriptive findings of the occasions when professors primarily made modifications to their own teaching practice, where feedback from students (90.9%), results from student evaluations (74.9%), and collegial exchange (67.5%) are among the top three. However, in terms of didactic professional development, the professors in our sample have rather settled down into some kind of "maintenance mode regarding teaching" (Lumpkin, 2009, p. 203).

6   Concluding remarks

Strikingly, what distinguishes this last project from its predecessors is that it also included a significant proportion of qualitative research. In a mixed-methods research, a frequent argument for such an approach is that qualitative and quantitative methods each have their own strengths for answering specific research questions and that, therefore, both strands of methods can be combined to compensate for the other's weaknesses (Kelle, 2008, p. 15). On that note, the variety of methods in empirical social research can be largely seen "a collection of craft skills, driven by local and practical concerns" (Seale; 1999, p. 26); that is, a practical toolkit to answer concrete research questions and to achieve specific epistemic goals. The specific epistemic goal of the first phase was to find answers to the question of how and why German universities try to raise the status of academic teaching in addition to monetary incentives. To this end, the latent construct of transformational governance first had to be filled with empirical content. For this purpose, the reasons and motives had to be understood first—why is it worthwhile for universities to give teaching an above-average status? Or, in the words of Schutz (1962):

> But the observational field of the social scientist—social reality—has a specific meaning and relevance structure for the human beings living, acting, and thinking within it. By a series of common-sense constructs they have pre-selected and pre-interpreted this world which they experience as the reality of their daily lives. It is these thought objects of theirs which determine their behavior by motivation it. The thought objects constructed by the social scientist, in order to grasp this social reality, have to be founded upon the thought objects constructed by the common-sense thinking of men, living their daily life within their social world. Thus, the constructs of the social sciences are, so to speak, constructs of the second degree, that is, constructs of the constructs made by the actors on the social scene, whose behavior the social scientist has to observe and to explain in accordance with the procedural rules of his science. (Schutz, 1962, p. 59)

To reconstruct these specific meanings and relevance structures in regards to the leading research question, a qualitative approach was necessary. The extreme case sampling of four teaching-awarded universities provided us with rich accounts from multiple perspectives on how these universities credibly try to increase the status of teaching at their institution. Semi-structured expert interviews seemed to be the most appropriate data collection strategy for this purpose.

The epistemological goal of the second phase was to determine whether and how the professors at the respective universities perceive these cultural and structural measures which necessitated a quantitative approach. Indeed, the comparison of teaching-awarded with not awarded universities in the follow-up survey confirmed our suspicion that the state of development of a supportive teaching culture is more advanced at teaching-awarded universities. However, at the level of action, a difference could only be observed at the universities of applied sciences; that is, professors at teaching-awarded universities of applied sciences are more committed to teaching than at not-awarded universities of applied sciences. Finally, the national survey allowed for empirically more reliable statements on the impact of different forms of teaching governance on individual teaching behavior.

The question of how much the research process, especially in the first phase, has followed the rigor and linearity of text book accounts on research designs and methods must be answered with 'more or less'. As Miles and Huberman (1994) put it,

> Some methodological texts tend towards the abstract, with brief examples that always seem to work out clearly, even effortlessly. Yet when you actually come to grips with collecting and analyzing real-life data, things seldom work out that way. Research-in-use is almost always more intractable, disjointed, and perverse than research-in-theory [...] In short, doing qualitative analysis is the way you get better at it—and we believe that holds true not just for novices and new entrants to qualitative work, but for wise old dogs as well. (p. 309)

In particular, the rich stock of case study research literature that I have dealt with in the context of this methodological reflection has not been very helpful in retrospect to justifying the research procedure. Actually, Yin's (2003) much quoted definition that "case studies are the preferred strategy when 'how' or 'why' questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context" (p. 1), hits the nail on the head, which constituted the multiple-case study for this particular research project, but other textbooks on qualitative research methods had to be consulted for our practical implementation and data analysis. In contrast, the review of the methodology of the expert interview with regard to the results achieved in the first phase was particularly fruitful and helped us to reflect the dominant data collection method in the retrospective. For the expert interviews to be successful, the interviewees had to be prepared to reveal their knowledge of interpretation and to engage in a conversation at

eye level. This was (fortunately) guaranteed thanks to the purposive sample of teaching-awarded universities that allowed us to only recruit those interview partners who, due to their supposed interest in our research project, would almost certainly reveal their specialized knowledge. This was ensured by the special interviewer constellation because both the project leader and an already experienced junior academic set the tone in the interviews.

In the retrospective, the question of what would be done differently in a re-design of the research project cannot be answered in general terms. In addition to interviews, it would have been conceivable to systematically incorporate more ethnographic elements in the form of participating observations of teaching-related events (such as on the Day of Teaching) at more universities. With regard to an empirically-based type formation (Kelle & Kluge, 1999) of the German professorship in regards to teaching, the exploration of the motives and relevance structures of further interviewees who rather represent the typical mainstream professor would also have been crucial. Finally, too little attention was paid in the quantitative surveys to the research aspect of being a professor. Although the multiple-case study enabled us to identify the conditions under which and why professors became more involved in teaching, we found that there was a clear pattern where these professors were not exclusively socialized in academia but held senior management positions in the private sector for a considerable amount of time. Accordingly, their academic careers were less dependent on research achievements.

In terms of the findings, the qualitative phase was probably the most revealing. Although we have largely dispensed with terms such as causality and generalizations of any kind, this does not detract from the value of the results. It was exactly the strength of the qualitative phase to work out the peculiarities of each university and, as we wrote in Lauer and Schmid (2016), "to understand and learn from each university's first-hand experiences about their journey to and institutional continuation of 'teaching excellence'" (p. 114). The investigation of the four best-practice universities showed that it is exactly the intrinsic teaching motivation that must continue to be promoted and nurtured, not only in a transformational way from the top-down but also from the bottom-up. Thus, upgrading the status of academic teaching does not take place in a vacuum but depends on the interaction of many actors at different levels within the

university. Therefore, it is not really surprising that the rectorates have all remained somewhat guarded over how teaching can be managed in concrete terms while referring to the organizational particularities (such as high trust or cooperative management structures) of their own university. In an extreme case, we were able to witness that this can also start out as an exclusively bottom-up initiative, such as can be seen in the 'dynamic manager duo' in Schmid and Lauer (2016); thus, "what may appear as 'transformational' is an outcome of subordinates being 'high-performing'" (Alvesson & Kärreman, 2016, p. 10).

An international comparison showed how different higher education systems affect the steering capacity of academic teaching, which either necessitates a more transactional governance approach or a more transformational governance approach (Wilkesmann, 2016). Thus, for the concrete case of Germany, which is characterized as a "soft governmental regime" (Wilkesmann, 2016), it remains in the responsibility of those in power within the field of higher education to change the rules of the "only research counts" (Schmid & Lauer, 2016) game insofar that it also adequately takes teaching into account. Consequently, it will be exciting to find which new funding programs will be launched when the 'Quality Pact Teaching' expires in 2020. Furthermore, private foundations such as the Stifterverband will continue to play an important role in the advancement of the status of teaching within the German higher education landscape. Finally, developments such as the digitalization of teaching and the expansion of university-wide quality management systems will likewise impact the 'how' and 'why' of academic teaching and learning at German universities in the long run.

The question of 'why do professors teach how they teach?' was already posed in Wilkesmann and Schmid (2011) and theoretically elaborated by Schmid (2016). For the case of professors at research universities, a short answer to this question is because teaching is mainly intrinsically motivated where teaching as a collectively valued practice is still in its infancy (Lauer & Wilkesmann, 2019). As one vice-rector at a research university states:

> First, own motivation plays an essential role to be committed in academic teaching. As one of my colleagues always says 'teaching is a private affair'; when you don't bring the passion for teaching with you, then to get motivated all at once through incentive systems is a tough undertaking. (research university, vice-rector for teaching and learning)

Hence, it is not that surprising that the quantitative results from the nationwide survey have only identified a few further factors on the university level that significantly contribute to the amount of explained variance in individual teaching behavior. Nonetheless, a university can actively signal in the form of transformational governance that a commitment to quality teaching is valued and appreciated by providing the leeway and autonomy that nurtures more internalized modes of teaching motivation. In the best case, this can lead to university-wide innovations, as shown in Lauer and Wilkesmann (2017). At the same time, transactional governance is necessary to reproduce and safeguard organizational routines but also to activate more externalized modes of teaching motivation that such newly introduced teaching innovations are enforced university-wide. However, longitudinal data is required to investigate if organizational change or learning has really occurred.

Besides replicating the strong influence of intrinsic teaching motivation, gender, and discipline, it is of note that the immediate teaching environment gained influence on individual teaching behavior. Nonetheless, future research should focus more on departmental teaching cultures to examine these dynamics in more depth. A no less interesting question would be how professors actually manage teaching at their own chairs. Here, for example, a survey of junior academic staff could be carried out to assess their motivation to teach and perception of their local teaching climate. Ideally, this would take concrete shape, as one professor at one of the teaching-awarded research universities puts it:

> So I am already thinking a lot, a lot about our teaching with my staff. And this is by no means transactional, I hope I interpret it correctly now, because I cannot offer them anything in return. I can't offer a salary increase or a faster option to complete their PhDs, on the contrary probably, it will take longer. And yet, as I have just said, we succeed in getting people to almost hit each other's heads because they absolutely want to teach, even those who actually don't have to. They want to do it simply because they enjoy it, because they do it well and it gives everyone a pleasure. And I think, if I have understood it correctly, it is just such an element of transformational leadership. (research university, professor)

7   Literature

Alvesson, M., & Kärreman, D. (2016). Intellectual failure and ideological success in organization studies: The case of transformational leadership. *Journal of Management Inquiry, 25(2)*, 139–152.

Argyris, C., & Schön, D. (1978). *Organizational Learning: A Theory of Action Perspective*, Reading, MA: Addison Wesley.

Arrow, K. J. (1985). The Economics of Agency. In J. W. Pratt, R. J. Zeckhauser & J. J. Arrow (Eds.), *Principals and Agents: The Structure of Business* (pp. 37–51). Boston, MA: Harvard Business Press.

Bass, B. M., & Avolio, B. J. (1993). Transformational leadership and organizational culture, *Public Administration Quarterly, 17(1)*, 112–121.

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership*. New York, NY: Psychology Press.

Blumer, H. (1954). What is wrong with social theory?. *American Sociological Review, 19(1)*, 3–10.

Bogner, A., Littig, B., & Menz, W. (2014). *Interviews mit Experten: eine praxisorientierte Einführung*. Wiesbaden: Springer.

Bogner, A., & Menz, W. (2009). The theory-generating expert interview: epistemological interest, forms of knowledge, interaction. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing experts* (pp. 43–80). Basingstoke: Palgrave Macmillan.

Brockerhoff, L., Stensaker, B., & Huisman, J. (2014). Prescriptions and perceptions of teaching excellence: a study of the national "Wettbewerb Exzellente Lehre" initiative in Germany. *Quality in Higher Education, 20 (3)*, 235–54.

Burzan, N. (2016). *Methodenplurale Forschung: Chancen und Probleme von Mixed Methods*. Weinheim: Beltz Juventa.

Cashin, W. E. (1989). *Defining and evaluating college teaching (IDEA Paper No. 21)*. Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications.

Creswell J. W., Plano Clark V. L., Gutmann M. L., & Hanson W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioural research* (pp. 209–240). Thousand Oaks, CA: Sage Publications.

De Boer, H., Enders, J., & Schimank, U. (2007). On the way towards new public management? The governance of university systems in England, the Netherlands, Austria, and Germany. In D. Jansen (Ed.), *New forms of governance in research organizations: Disciplinary approaches, interfaces and integration* (pp. 137–152). Dordrecht: Springer.

Dee, J.R., & Leišytė, L. (2016), Organizational learning in higher education institutions: theories, frameworks, and a potential research agenda. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research Vol. 31* (pp. 275–348). Dordrecht: Springer.

Denzin, Norman K. (1989). *The Research Act: A Theoretical Introduction to Sociological Methods*. New York, NY: McGraw-Hill.

Deutscher Hochschulverband (2016). 2016/Band 1 Universitäten Deutschland. Berlin, Boston: De Gruyter Saur. Retrieved from https://www.degruyter.com/view/product/465994

DiMaggio, P. J. (1988). Interest and agency in institutional theory. In L. G. Zucker (Ed.), *Institutional Patterns and Organizations: Culture and Environment* (pp. 3–21). Cambridge, MA: Ballinger.

Eisenhardt, K. M. (1989a). Agency Theory: An Assessment and Review. *Academy of Management Review 14(1)*, 57–74.

Eisenhardt, K. M. (1989b). Building theories from case study research. *Academy of Management Review*, *14(4)*, 532–550.

Emirbayer, M., & Mische, A. (1998) .What is agency?. *American Journal of Sociology, 103(4)*, 962–1023.

Engel, U., & Schmidt, B. O. (2019). Unit- und Item-Nonresponse. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS.

Feldman, K. A., & Paulsen, M. B. (1999). Faculty motivation: The role of a supportive teaching culture. *New Directions for Teaching and Learning, 78*, 69–78.

Fernet, C., Senécal, C., Guay, F., Marsh, H., & Dowson, M. (2008). The work tasks motivation scale for teachers (WTMST). *Journal of Career assessment, 16(2)*, 256–279.

Flick, U. (2004). Triangulation in qualitative research. In U. Flick, E. von Kardoff & I. Steinke (Eds.), *A companion to qualitative research* (pp. 178–183). London: Sage Publications.

Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Harrison, H., Birks, M., Franklin, R., & Mills, J. (2017). Case Study Research: Foundations and Methodological Orientations [34 paragraphs*]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 18(1), Art. 19*, Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs1701195

Hartley, J. (2004). Case study research. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research* (pp. 323–333). London: Sage Publications.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15(9)*, 1277–1288.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling, 6(1)*, 1–55.

Hüther, O., & Krücken, G. (2018). *Higher Education in Germany: Recent Developments in an International Perspective*. Berlin: Springer.

Hunter, A., & Brewer, J. D. (2015). Designing Multimethod Research. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry* (pp. 186–205). Oxford: Oxford University Press.

Jorzick, B. (Ed.) (2013). *Charta guter Lehre. Grundsätze und Leitlinien für eine bessere Lehrkultur*. Essen: Edition Stifterverband. Retrieved from: https://www.stifterverband.org/download/file/fid/187

Kelle, U. (2008). *Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kelle, U., & Kluge, S. (1999). *Vom Einzelfall zum Typus*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kultusministerkonferenz (KMK) (2008), *Wettbewerb exzellente Lehre – Eine gemeinsame Initiative der Kultusministerkonferenz und des Stifterverbands für die Deutsche Wissenschaft [Competition for teaching excellent – a joint initiative of the KMK and the Stifterverband]*. Retrieved from http://www.exzellente-lehre.de/pdf/kmk_beschluss_06_03_2008.pdf

Land, R., & Gordon, G. (2015). *Teaching excellence initiatives: modalities and operational factors. York: Higher Education Academy*. Retrieved from https://www.heacademy.ac.uk/system/files/resources/teaching_excellence_initiatives_report_land_gordon.pdf

Lauer, S., & Wilkesmann, U. (2017). The governance of organizational learning – empirical evidence from best-practice universities in Germany. *The Learning Organization, 24(5)*, 266–277.

Lauer, S., & Wilkesmann, U. (2019). How the institutional environment affects collegial exchange about teaching at German research universities: findings from a nationwide survey. *Tertiary Education and Management, 25(2)*, 131–144. https://doi.org/10.1007/s11233-019-09020-5

Lumpkin, A. (2011). A model for mentoring university faculty. *The Educational Forum, 75(4)*, 357–368.

Mayring P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Klagenfurt: Monograph. Retrieved from http://nbnresolving. de/urn:nbn:de:0168-ssoar-395173

Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz Juventa.

Menold, N., & Züll, C. (2011). Standardisierte Erfassung von Verweigerungsgründen in Face-to-Face-Umfragen. *GESIS Methoden‑Daten‑Analysen 5 (1)*, 91–108.

Meuser, M., & Nagel, U. (2005). ExpertInneninterviews – vielfach erprobt, wenig bedacht: ein Beitrag zur qualitativen Methodendiskussion. In A. Bogner, B. Littig & W. Menz, *Das Experteninterivew* (pp. 71–93). Wiesbaden: VS Verlag für Sozialwissenschaften.

Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In A. Bogner, B. Littig & W. Menz (Eds.), *Interviewing experts* (pp. 17–42). Basingstoke: Palgrave Macmillan.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie, *Handbook of mixed methods in social & behavioral sciences* (pp. 189–208). Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*. Thousand Oaks, CA: Sage Publications.

Pfadenhauer, M. (2009). At eye level: The expert interview—A talk between expert and quasi-expert. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing experts* (pp. 81–97). Basingstoke: Palgrave Macmillan.

Pflüger, J., Pongratz, H. J., & Trinczek, R. (2010). Fallstudien in der deutschen Arbeits- und Industriesoziologie. In H. J. Pongratz & R. Trinczek (Eds.), *Industriesoziologische Fallstudien. Entwicklungspotenziale einer Forschungsstrategie* (pp. 23–70). Berlin: Edition Sigma.

Pflüger, J., Pongratz H. J., & Trinczek, R. (2017). Fallstudien in der Organisationsforschung. In S. Liebig, W. Matiaske & S. Rosenbohm (Eds.), *Handbuch Empirische Organisationsforschung* (pp. 389–413). Wiesbaden: Springer.

Ragin, C. C. (1992). Introduction: Cases of "What is a case?" In C. C. Ragin & H. S. Becker (Eds.), *What is a case?* (pp. 1–18). Cambridge, UK: Cambridge University Press.

Ramm, M. (2014*). Response, Stichprobe und Repräsentativität. Zwei Dokumentationen zum Deutschen Studierendensurvey (DSS)*. In Universität Konstanz Arbeitsgruppe Hochschulforschung, Fachbereich Geschichte und Soziologie (Series Ed.) Hefte zur Bildungs- und Hochschulforschung. Retrieved from https://kops.uni-konstanz.de/bitstream/handle/ 123456789/29793/Ramm_0-262244.pdf?sequence=3

Ridder, H. G. (2017). The theory contribution of case study research designs. *Business Research, 10(2)*, 281-305.

Rindermann, H. (2000). Das Selbstobjektivierungsproblem im akademischen Milieu (II): Zusammenhang zwischen wissenschaftlicher Methodik und dem Einfluss sozialer Hintergrundvariablen in der Lehrevaluationsforschung. *Das Hochschulwesen, 48(4)*, 117–123.

Robson, C., & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings*. Hoboken, NY: Wiley.

Ryan, R.M., & Deci, E.L. (2000a), Intrinsic and extrinsic motivation: classic definitions and new directions, *Contemporary Educational Psychology, 25(1)*, 54–67.

Ryan, R.M., & Deci, E.L. (2000b), Self-determination theory and the facilitation of intrinsic motivation, social development and well-being, *American Psychologist, 55(1)*, 68–78.

Schmid, C. J. (2016). *Die soziale Organisiertheit und Organisierbarkeit von Interessen(freiheit) – Der Fall der managerialen Governance akademischer Lehrtätigkeit*. (Doctoral dissertation). Retrieved from https://eldorado.tu-dortmund.de/bitstream/2003/34952/1/Dissertation_Schmid_25.04.2016_Epub.pdf

Schmid, C. J., & Wilkesmann, U. (2015). Ansichtssache Managerialismus an deutschen Hochschulen–ein empirisches Stimmungsbild und Erklärungen. *Beiträge zur Hochschulforschung, 37(2)*, 56–87.

Schmid, C. J., & Lauer, S. (2016). Institutional (teaching) entrepreneurs wanted! Considerations on the professoriate's agentic potency to enhance academic teaching in Germany. In L. Leisyte & U. Wilkesmann (Eds.), *Organizing academic work in higher education: Teaching, learning and identities* (pp. 109–131). London: Routledge

Schwandt, T. A., & Gates, E. F. (2018). Case study methodology. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed.) (pp. 341–358). Thousand Oaks, CA: Sage Publications.

Schutz, A. (1962). Concept and theory formation in the social sciences. In M. Natanson (Ed.), *Collected Papers, Vol. 1, The Problem of Social Reality* (pp. 48–67). The Hague: Martinus Nijhoff.

Seale, C. (1999). *The Quality of Qualitative Research*. London: Sage Publications.

Sjoberg, G., Williams, N., Vaughan, T., & Sjoberg, A. (1991). The case study approach in social research. In J. R. Feagin, A. M. Orum & G. Sjoberg, *A case for the case study* (pp. 27–79). Chapel Hill, NC: University of North Carolina Press.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel Analysis*. Thousand Oaks, CA: Sage Publications.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.) (pp. 443–466). Thousand Oaks, CA: Sage Publications.

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: The Guilford Press.

Stifterverband (2019). Genius Loci-Preis für Lehrexzellenz [Genius Loci-Award for Teaching Excellence] Retrieved from https://www.stifterverband.org/genius-loci

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage Publications.

Teddlie, C., & Yu, F. (2007). Mixed Methods Sampling. *Journal of Mixed Methods Research, 1(1)*, 77–100.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage Publications.

Tight, M. (2010). The curious case of case study: a viewpoint. *International Journal of Social Research Methodology, 13(4)*, 329–339.

Tobin, R. (2010). Quick start to case study research. In A. J. Mills, G. Durepos & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp. 771–773). Thousand Oaks, CA: Sage Publications.

Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse, and structure. *Qualitative inquiry, 17(6)*, 511–521.

Trigwell, K., Prosser M., & Waterhouse, F. (1999). Relations Between Teachers' Approaches to Teaching and Students' Approaches to Learning. *Higher Education 37 (1)*, 73–83.

Weinhardt, M., & Liebig, S. (2015). Teilnahmeverhalten und Stichprobenverzerrung in der deutschen Stichprobe des European Social Survey. In J. Schupp & C. Wolf (Eds.), *Nonresponse Bias* (pp. 47–83). Wiesbaden: Springer.

Wilkesmann, U. (2013). Effects of transactional and transformational governance on academic teaching: Empirical evidence from two types of higher education institutions. *Tertiary Education and Management, 19(4)*, 281–300.

Wilkesmann, U. (2015). Transformationale Governance von Lehrkulturen an Hochschulen–Empirische Evidenz von zwei Hochschultypen in Deutschland. In M. Fischer, I. Langemeyer, & M. Pfadenhauer (Eds.), *Epistemic and Learning Cultures: Wohin sich Universitäten entwickeln* (pp. 96–114). Weinheim: Beltz Juventa.

Wilkesmann, U. (2016). Teaching matters, too: Different ways of governing a disregarded institution. In L. Leisyte & U. Wilkesmann (Eds.), *Organizing academic work in higher education: Teaching, learning and identities* (pp. 33–54). London: Routledge.

Wilkesmann, U. (2017). Methoden und Daten zur Erforschung spezieller Organisationen: Hochschulen. In S. Liebig, M. Wenzel & S. Rosenbohm (Eds.), *Handbuch Empirische Organisationsforschung* (pp. 565–588). Wiesbaden: Springer.

Wilkesmann, U. (2019). *Methoden der Hochschulforschung. Eine methodische, erkenntnistheoretisch- und organisationstheoretische Einführung*. Weinheim: Beltz Juventa.

Wilkesmann, U., & Lauer, S. (2015). What affects the teaching style of German professors? Evidence from two nationwide surveys. *Zeitschrift für Erziehungswissenschaft, 18(4)*, 713–736.

Wilkesmann U., & Lauer, S. (2018). The influence of teaching motivation and new public management on academic teaching. *Studies in Higher Education*. Advance online publication. https://doi.org/10.1080/03075079.2018.1539960

Wilkesmann, U., & Schmid, C. J. (2011). Lehren lohnt sich (nicht)? Ergebnisse einer deutschlandweiten Erhebung zu den Auswirkungen leistungsorientierter Steuerung auf die universitäre Lehrtätigkeit. *Soziale Welt 62*, 251–278.

Wilkesmann, U., & Schmid, C. J. (2012). The impacts of new governance on teaching at German universities. Findings from a national survey. *Higher Education, 63(1)*, 33–52.

Wilkesmann, U., & C. J. Schmid. (2014). Intrinsic and Internalized Modes of Teaching Motivation. *Evidence-based HRM: A Global Forum for Empirical Scholarship 2 (1)*, 6–27.

Yazan, B. (2015). Three approaches to case study methods in education: Yin, Merriam, and Stake. *The qualitative report, 20(2)*, 134–152.

Yin, R. K. (2003). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

8   Appendix

Appendix 1. Confirmatory factor analysis for teaching motivation

Standardized regression weights	Items
Intrinsic teaching motivation (Cronbach's Alpha=.71)	
.79	Because I derive much pleasure for my teaching.
.81	Because I find teaching interesting.
.55	Because I lose myself in teaching.
Identified teaching motivation (Cronbach's Alpha=.72)	
.76	Because I see my teaching as a significant contribution to my students overall academic progress.
.80	Because for me, the task of teaching is of personal importance.
.62	Because the task of teaching provides the chance to realize an aspect of my academic profession that is of personal meaning to me.
Introjected teaching motivation (Cronbach's Alpha=.61)	
.76	Because I have a bad conscience if I neglected my teaching duties.
.37	Because a good performance in teaching attributes largely to my self-esteem as a professor.
.62	Because I would feel bad if I neglected my teaching duties.
Extrinsic teaching motivation (Cronbach's Alpha=.63)	
1.00	Because my university/employment contract demands to teach.
.46	Because I get paid for it.
Amotivation (Cronbach's Alpha=.66)	
.55	I don't know why, because the work conditions provided for academic teaching are unbearable.
.77	I don't know why, sometimes I don't see the actual purpose of teaching.
.74	I don't care much for teaching as I don't know what it effects.

Appendix 2. Correlations of the latent variables

	Intrinsic teaching motivation	Identified teaching motivation	Introjected teaching motivation	Extrinsic teaching motivation	Amotivation
Intrinsic teaching motivation		.66**	.17**	-.23**	-.47**
Identified teaching motivation			.12*	-.17**	-.46**
Introjected teaching motivation				.24**	-.05
Extrinsic teaching motivation					.19**

** sig. 0.01, * sig. 0.05

9   List of publications contributing to the cumulative dissertation

In the following, all four contributions to this cumulative dissertation are listed in chronological order:

Article 1:

Schmid, C. J., & Lauer, S. (2016). Institutional (teaching) entrepreneurs wanted! Considerations on the professoriate's agentic potency to enhance academic teaching in Germany. In L. Leisyte & U. Wilkesmann (Eds.), *Organizing academic work in higher education: Teaching, learning and identities* (pp. 109–131). London: Routledge. https://doi.org/10.4324/9781315693729

Article 2:

Lauer, S., & Wilkesmann, U. (2017). The governance of organizational learning - empirical evidence from best-practice universities in Germany. *The Learning Organization, 24(5)*, 266–277. https://doi.org/10.1108/TLO-02-2017-0012

Article 3:

Wilkesmann U., & Lauer, S. (2018). The influence of teaching motivation and new public management on academic teaching. *Studies in Higher Education*. Advance online publication. https://doi.org/10.1080/03075079.2018.1539960

Article 4:

Lauer, S., & Wilkesmann, U. (2019). How the institutional environment affects collegial exchange about teaching at German research universities: findings from a nationwide survey. *Tertiary Education and Management, 25(2)*, 131–144. https://doi.org/10.1007/s11233-019-09020-5