



Citation for published version:

Pineda, P & Tim, S 2022, 'Translating student evaluation of teaching: how discourse and cultural environments pressure rationalizing procedures', *Studies in Higher Education*, vol. 47, no. 7, pp. 1326-1342.
<https://doi.org/10.1080/03075079.2021.1889491>

DOI:

[10.1080/03075079.2021.1889491](https://doi.org/10.1080/03075079.2021.1889491)

Publication date:

2022

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY-NC

Accepted Author Manuscript. This is an Accepted Manuscript of an article published by Taylor & Francis in *Studies in Higher Education* on 16 Feb 2021, available online:
<http://www.tandfonline.com/10.1080/03075079.2021.1889491>

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Translating Student Evaluation of Teaching: How Discourse and Cultural Environments Pressure Rationalizing Procedures

Wordcount: 9032 words

Pedro Pineda ORCID: [0000-0001-8775-1971](https://orcid.org/0000-0001-8775-1971)^{a b} and Tim Seidenschnur ORCID [0000-0003-4688-4182](https://orcid.org/0000-0003-4688-4182)^c

^a First affiliation: German Centre for Higher Education Research (DZHW), Hannover, Germany

^b Second affiliation and ^b First affiliation: International Centre for Higher Education Research Kassel (INCHER), University of Kassel, Germany

Funding Sources

This work was supported by the Fritz Thyssen Foundation through Research Scholarship to Pedro Pineda [number Az. 40.18.0.003SO].

Acknowledgements

We would like to thank Francisco Ramirez for all his support during the preparation and fieldwork of this investigation. We also would like to express gratitude to Katherine Bird for proofreading this work.

Translating Student Evaluation of Teaching: How Discourse and Cultural Environments Pressure Rationalizing Procedures

Student evaluation of teaching (SET) has not yet been studied historically and comparatively. Based on our interviews with professors and administrators at 18 universities in three countries, we discuss how SET diffused in all the studied universities and how SET was translated and edited differently in these countries according to different sets of statements. SET diffused from the US, where it was initiated by students in the 1950s and later adopted by universities in the 1970s. German and Colombian universities only imported SET later, in the 2000s, and in the German public sector it is still not mandatory for all courses. SET gains legitimacy through different discourses, for instance, the discourse on the empowerment and rights of students. SET also connects to the discourse on increasing rationality that extends the metrification of teaching to universities and is often connected to discourses on competition. But SET has also been reformed because it is regarded as biased against underrepresented groups and faces criticism of its methodological validity. By analysing differences and similarities in the legitimizing discourses and critiques across countries and sectors, we aim to contribute to the theorization of diffusion, translation and editing of organizational practices. In the case of evaluation practices, discourses on rationalization, student empowerment and competition are used to gain legitimacy while academic discourses on academic freedom and social diversity discourses are used to oppose SET, make translation more difficult and lead to editing of SET practices.

Keywords: student evaluation of teaching; teaching evaluation; quality assurance; neo-institutionalism; social change; academic discourse; organizational change

Introduction

Neo-institutional theory has been applied to study procedures that are comparable to student evaluation of teaching (SET), such as rankings or ratings of scientific production (Meyer and Ramirez 2013). However, the diffusion proposition has not yet been applied to the study of SET beyond national borders. If SET were rejected or substantially different in universities from different traditions, then alternative local (Wedlin and Sahlin 2017) or ‘glocalization’ (Drori, Höllerer, and Walgenbach 2014) explanations could prove more suitable to explain the transfer of ideas and practices in higher education. In relation to this homogenization debate, a broader analysis addressing the underlying discourses that have favoured or hindered the acceptance of arguments for or against SET’s implementation also sheds light on how the global movement of SET takes place according to different sets of statements in different national academic systems. Studying SET under a comparative perspective thereby contributes to theorizing diffusion, translation and editing processes in new institutional theory and in organization studies in particular.

To investigate if, why and how students from universities within different national systems start to evaluate teaching, we visited 18 universities in the United States (US), Germany and Colombia. Our inquiry overlaps with previous neo-institutional research that found a global tendency among universities in different places to rationalize their activities (Meyer and Ramirez 2013; Ramirez, Meyer, and Lerch 2016), but we also point out that other sets of ideas become relevant for the translation of SET into different local systems during institutionalization: the empowerment of students, professorial autonomy, and concerns for diversity. We thus analyse if and how the different discourses become relevant for the translation of SET in different countries.

We first introduce the debate between a neo-institutional theoretical perspective and alternative approaches in organizational studies to understand the role of the university in terms of an institutional locus for encountering global practices or a local organization influenced by imported ideas. We emphasize the mediating role of social discourses to further theorize translation and editing during the transfer of ideas. Afterwards, we describe our comparative approach, which allows us to present the statements for or against implementing SET in the universities from each country. After discussing similarities and differences across national settings and sectors, we use the analysis of SET to further theorize the transfer of evaluation practices as being representative of the global diffusion of organizational practices. Based on the social discourses we discovered, we propose that globalization and localization of SET can only be understood by examining the interplay of discourses compatible with or contradictory to evaluation practices.

Theoretical Framework

The diffusion hypothesis in neo-institutional theory refers to the worldwide spread of practices and in particular those that a rationalistic world culture promotes (Meyer et al. 1997). According to DiMaggio and Powell (1983), increasing rationalization has the paradoxical effect of creating similar forms of organization, structures and practices such as SET. Organizations become increasingly similar as their environments share similar cultural determinants pressuring for efficiency seeking in a similar task environment (Scott 1995; Boxenbaum and Jonsson 2017). The increasing transfer practices that standardize administration with a managerial character have been encountered in business sectors, but also in non-profit, government, and higher education sectors (see Bromley and Meyer

2014). With increasing social interconnection, rationalization was first a Western process but since the Second World War an increasingly global trend.

Systems of rationalization are preferred over traditional practices because they acquire greater legitimacy in their imaginary connection to human progress. This rationalization of social life also tends to be a global trend because both nation states and sectors of social life have become increasingly interconnected after the establishment of a new form of global governance after the Second World War (Krücken and Meier 2006). Therefore, the measurement of activities in rationalized terms to attain goals is seen as a universal rule that not only diffuses across national states but also through different types of organizations, including universities. If successfully globalized, the diffusion of SET would be an example that proves validity of the core neo-institutional proposition of diffusion.

The study of the global diffusion of management trends and models may also consider the local dimension. Drori, Höllerer, and Walgenbach (2014) revitalize the importance of the local dimension in comparative studies but overcome dichotomies between the global and the local. Empirical research, they propose, should focus on ideas, practices and structures that are transferred, the actors and the mechanisms that propel the process, considering temporal and interrelated places—from the local to the global. They reframe the term ‘glocalization’ as it appeared in the 1990s to describe the global influence of the Western world (Wellman and Hampton 1999) to the study of the import of managerial practices at the level of organizations in order to emphasize “dualities of similarity and variation as well as universalism and particularism” (Drori, Höllerer, and Walgenbach 2014, 86). Also influenced by this approach, our simultaneous study of SET in different contexts – where it was created and where it was later adopted – can also highlight the global and local components during the transfer of an archetypical evaluation practice.

Still, the role of discourses in transforming ideas from abroad tends to be underestimated in the literature about organizations and in higher education scholarship. Management models are not only transferred but they are also translated and edited. They are linked to discourses already accepted as appropriate to the local settings into which they are imported (Wedlin and Sahlin 2017; Author 1 2020). Such settings are usually heterogenous and so are the discourses in different institutional settings that may be used to supplement or question the appropriateness of the imported model. Actors legitimize and justify these imported rationalized practices into local systems (Drori, Höllerer, and Walgenbach 2014). Since institutional environments differ, the idea or practice must not only be translated differently but is also transformed or edited in a different manner in different organizational contexts (Sahlin-Andersson 1996; Wedlin and Sahlin 2017; Waeraas and Nielsen 2016). New management practices “get translated throughout their circulation, and as they evolve differently in different settings, they do not only lead to homogenization but also to variation and stratification” (Wedlin and Sahlin 2017, 103). The whole process of appropriation of external ideas can lead to different ecologies of types of the originally diffused model (Greenwood et al. 2011; Kraatz and Block 2008). Regional or sectorial similarities and differences are also plausible during translation and editing processes.

We understand discourses in line with Foucault: As a formation of knowledge that determines how actors perceive, interpret and accept instruments put into operation. Discourses are evidenced when actors reproduce their elements of meaning, symbols, and interpretative patterns (Foucault 1969/2002). Discourses are built up by groups of statements that actors articulate to ascribe meaning to what they say, think, and do. In higher education, discourses in favour of rationalization practices are used to legitimize

increasing accountability in universities. For example, new public management (NPM) doctrines applied to higher education promote the understanding that universities must continuously prove that they are establishing and reaching goals in their different activities in order to maximize their resources (Kwak, Gavrilu, and Ramirez 2019; Krücken and Meier 2006). Discourses emphasizing competition may partially conflate with debates on increasing rationalization. In the case of SET, competition is evidenced if the results are used not only to provide feedback to the lecturers but also to feed into reports for gaining material or symbolic resources (e.g. bonuses or teaching awards) (see Krücken 2019). An economically-oriented rationale may reproduce discourses of SET that involve students in their role as clients paying for a teaching service (Garcia Garduno 2014; Sproule 2000). Emancipatory discourses may describe SET in terms of an instrument that protects student rights, giving them the opportunity to control teaching activities of the universities (Lerch et al. 2016). The study of SET is thus exemplary to empirically study the role of enabling and contradicting discourses within the broader debate of diffusion or glocalization of ideas and practices in the exemplary context of higher education.

Method

We chose to compare different cases from the US, Germany, and Colombia, which represent different higher education traditions, to test the diffusion and translation propositions in SET. We selected 18 universities that represented typical cases for each country (Ebbinghaus 2016; Lijphart 1971). We included three public universities and three private universities from each country out of the total number of 418, 103 and 58 higher education institutions officially classified as universities and capable of awarding doctorates (SNIES 2020; IPEDS 2020; ICEland 2020). The private sector is large and even

majoritarian in terms of number of institutions in the US, which has 187 (17 of them for-profit), and in Colombia, which has 47, while in Germany only 20 universities are private. Still, we included German private universities to maintain the cross-national comparison across sectors. We considered private universities because they could be places for legitimizing accountability procedures—including teaching—in Germany, also because of the role that tuition fees (prohibited in the public sector, see Hüther and Krücken 2014) may play in changing status and power of actors involved. We could not compare for-profit universities cross-nationally because higher education must be nonprofit in Colombia and Germany.

We chose universities from US, Germany and Colombia to include the national level, which is where some social discourses surrounding the implementation of SET might differ. Our selected countries have experienced different path dependencies and represent higher education traditions in the US, continental Europe and Latin America (Krücken 2003; Espeland and Sauder 2007; Brittingham 2009; Author 1 2015). Still, there was at least some degree of discussion on the implementation of SET in each of these settings (Wolbring 2010; Hildebrand, Wilson, and Dienst 1971; Villavicencio 2008; Rindermann 1995).

Higher education in the US is characterized by greater mechanisms of accountability (Ramirez 2010). Emphasis on productivity is related to less power for the professoriate in relation to other countries (Ben-David and Collins 1966). German public universities, accounting for 92% of university enrolment (ICEland 2020), are still highly reliant on public funds and, more recently, competitive public funds (Krücken 2019) but not tuition fees. Hence, student empowerment driven by their participation in funding the university could be a more minor force, although other forms of legitimacy to participate in

academic bodies are well established. Colombian higher education, as a representative of Latin American higher education, is influenced by the different European traditions on different scales and over varying periods of time: the Spanish and French traditions, followed by the Humboldtian model, later local developments from the Córdoba reform and, most recently the global influence of the entrepreneurial model from the US (Author 1 2015). The politicized character of Latin American universities inspired by the Córdoba reform also reduced the influence of government intervention but was mixed with an ideal of democratization that in some cases even allowed both students and lecturers to vote for rectors and academic leaders under the principle of *co-gobierno* (Author 1 2015). These different traditions made us expect that different discourses are present in each country. These settings may combine transnational discourses that diffuse with more traditional ones.

We visited all universities and interviewed officers or university professors in charge of SET. These focused interviews (see Merton and Kendall 1946; Cohen, Manion, and Morrison 2018) were conducted with 13, 15 and 14 interviewees in the US, Germany, and Colombia. Interviewees were selected through consulting web pages of each university. The organizational unit in which they were located in the university varied greatly, and they could be linked to university registry, the vice-president in charge of teaching and learning, centers for teaching and learning, offices for quality assurance or within the faculties. We visited and interviewed at least two interview partners in each university. We also conducted document analysis and collected university regulations, manuals, evaluation forms and all the information available on the universities' web pages.

After collecting the data, we transcribed the interviews in MAXQDA, coded the responses and conducted a content analysis according to our categories of analysis (see

Annex 1). Given that “discourse is the conditions under which certain statements are considered to be the truth” (Ball 2013, 19), we aimed to understand the context of implementing SET in relation to the symbols, statements and groups of statements that constitute discourse (Foucault 1969/2002, 32). New categories emerged on the rationale to import, translate or edit SET.

Diffusion, Translation and Editing of Student Evaluation of Teaching

United States

Diffusion

All universities visited in the US applied SET for all courses. The five universities that had a starting date we could verify started to implement SET between 1975 and 1984 (see Table 1). SET administered by students based on informal questionnaires, though, was a precursory practice in the US, as is evidenced by the Scratch Sheet, a leaflet published by the Associated Students of Stanford University (ASSU) since at least 1958 (Associated Students of Stanford University 1958; Stanford Daily 1958).

[Table 1 Student evaluation of teaching in universities in the US]

Statements in public and private universities

Both the public and private universities we visited use SET among other sources of information mainly to take decisions on academic careers: hiring, promotion, tenure processes and pay, including using SET as a source of information in competitions for teaching awards (in AU1). At AU1, AU4 and AU6 evaluation results are made available to

students, and professor AU4.2 even recalls that the administration department used to publish the results on billboards. The use of SET for building aggregated indicators and monitoring is not central. Also, lecturers present their own SET scores when competing for new employment, although in the opinion of the person in charge of SET: “It is not really a legitimate [arithmetic] mean but that is what we would do”, explains professor AU3.1 as he shows the folder with SET results he has archived during his teaching career. SET administrators from AU1, AU3 and AU4 said that, for adjunct lecturers, scores were the main sources of information for deciding whether to employ them: “For adjunct they are perhaps even more important because the scores they get will decide if they will be retained or not, especially in the case where there is a lot of competition for these jobs” (AU3.1).

With regard to the role of students, students at AU1 and AU4 even achieved the power to request information when selecting new courses, regardless of the validity problems the instruments may have or the consequences for their relationship with the faculty in terms of the risk of isomorphic teaching practices generating grade inflation. Knowing that SET was initially a student initiative before universities took over its control allows us to understand that students may hinder major changes in its implementation.

Accreditation is said not to have caused the implementation of SET: “It was not spurred by accreditation, not in response to or specifically for accreditation. It was more for internal purposes that certainly help for accreditation later on but that has never been impetus” (AU1.4), says a professor from AU1. Program accreditation seems to occur as a parallel and not causal process to the establishment of SET, since it was institutionalized at the same pace since the Higher Education Act of 1965 and most recently controlled by the Council for Higher Education Accreditation (CHEA), founded in 1996 (Brittingham 2009). AU2 and AU3 used Class Climate Software by Scantron (an American company

specialized in administering educational tests and surveys) for analysing and creating reports but also used forms internally designed by academics. AU6 first used standardized instruments: the Hildebrand, Wilson, and Dienst instrument (Hildebrand, Wilson, and Dienst 1971), developed at Berkeley and UC Davis, but since 2015 has started to use a home-made instrument.

Translation and Editing of SET

All universities visited, apart from public AU3, have recently discussed wide-ranging reforms to SET. A strong reformist spirit can be found in public universities AU1 and AU2 where SET has started to be transformed, mainly based on critiques of its discriminatory effects on underrepresented groups. Therefore, it is recommended to triangulate the quantitative results of the surveys with other sources of information. Professor AU1.2, in his role as former department head and dean at AU1, in 2012 banned the use of averages and indicators, and changed the presentation of surveys results to dossiers and peer observations. His decisions were reversed at the faculty level because the new instruments were seen as too costly, but they were maintained in his department. Similar tensions were also present in different departments of AU3. From his perspective as director of the Teaching Center of the university, professor AU1.4 explains:

“Voices have become louder around the issues of student evaluations, we see faculty taking greater ownership, starting to compare elements of or full dossiers. Say ‘I am not going to leave this up to the review committee to just decide how good I am or how excellent I am teaching based on my evaluations. (...) I am going to build an arsenal of evidence to show my teaching effectiveness’” (AU1.4).

In line with this trend, AU2 also issued a new policy that requires at least two measures out of ten listed when measuring “excellence in teaching” for appointment, promotion or appraisal decisions. In January 2018 the university changed the name from “student satisfaction surveys” to “student experience of teaching surveys”; each lecturer is completely free to design the questionnaire for their class. The professor in charge of SET at AU2 explained: “we are moving away from this paradigm, these are not evaluations but students describing their experiences”, initiated by “a group of faculty who is driving the changes” because “this is a campus that cares about diversity, cares about social justice. As we are getting more research telling that evaluations tend to be biased, we don’t want, don’t want to make decisions based purely on biased information” (AU2.1). AU2.1, thus, did not question the institutionalized belief in the necessity for involving students in the evaluation of teaching but argued that his university emphasizes feedback over monitoring.

Existing discussions about discrimination arising from using SET have followed other courses at private universities. At AU6 an adjunct professor argued in 2011 that her department had a gender bias evidenced in results from SET on which they based strategic actions. The central administration, in turn, described these concerns as “beliefs or myths” held by the faculty. “The issue of gender bias was raised initially eight years ago” by a female professor of the (...) department who “tried to get momentum to raise issues about that and it just didn’t happen” (AU6.1), said the person in charge of SET at AU6. The request was responded to negatively by the administration, which on the basis of a larger study comparing similar courses argued that there was not enough evidence for gender bias. The university, however, did create a Task Force in 2014 on the Evaluation of Teaching, consulted the faculty and revised surveys elsewhere. The previous instrument was

shortened and adapted to better address questions directed at measuring “excellence in teaching” based on changes in educational theory but where the bias issue was not central.

A process of questionnaire optimization took place at AU6, where since 2015 it has been shortened to 13 questions “through the psychometrics and factorial confirmatory analysis” of questions that “reflected our values at AU6, that encourage students in teaching and learning” (AU6.1), explained a member of the university leadership. The faculty union, for its part, is requesting less reliance on students in evaluation and tenure promotion, while the university has offered confidential observations that can be added to the tenure record, but has difficulties in adopting them as a standardized process; “The administration wants to have the conversation with the union. (...) but we also need to have something that will be implemented consistently across the faculty”, explains AU6.2, the person in charge of SET. Those responsible at AU6 seem to be more concerned about interrater validity than bias toward underrepresented groups. At AU4, diversity debates were also not central. In 2014, a Course Evaluation Committee modified previous forms to a shorter, customizable questionnaire that reduced response errors from the surveys, which usually had the same answers to all the questions. But even after this process, the administrator of the process AU4.1 perceives a “passive resistance” from some tenured professors about the excessive weight they consider is attached to SET results in tenure processes.

Other universities outside our case selection such as the University of Colorado Boulder (Reinholz et al. 2018) and the University of Southern California (University of Southern California 2018) have also stopped focusing on quantitative indicators from SET surveys. More recently, the American Sociological Association (2019) published a statement urging the use of SET for providing student feedback (useful also if visualized as a trend over time) rather than for creating rankings of teaching effectiveness, especially if

used without other sources of information. Diversity and validity discourses also seem to have been adopted in the translation of SET in universities outside our sample.

Germany

Diffusion

The German universities we visited started to institutionalize SET between 2000 and 2013, but we can also observe an ongoing process. The academic discussion of SET began with the dissertation and publications of Heiner Rindermann at the University of Heidelberg (Rindermann 1995, 1997). When we started our visits, GU2 said that SET was not being applied in the faculty and an extensive interview was not granted because “The student teaching evaluations are still in their infancy. Therefore, I cannot tell you much about them and am not sure if I could help you” (GU2.3).

[Table 2 Student evaluation of teaching in German universities]

Statements in public and private universities

Both coercive and normative forces have influenced the adoption of SET in Germany in a top-down manner, starting from national and federal regulations and funding down to university procedures. Both GU1 and GU3 also reported that they have used government funding from the Quality Pact for Teaching (2011–2016 and 2016–2020) to establish SET. The projects financed with this money include doctoral dissertations and positions in charge of SET (which we found in all universities except the smaller GU5). The Pact is an agreement between central and federal governments that aims to “improve studying conditions and increase teaching quality” (BMBF 2018, 2) by means of projects

addressing 11 topics including “quality assurance” “teaching-learning concepts”, and “valuation of teaching”. According to the project database (BMBF 2019), 18 universities have received funding for such projects.

In both public and private universities SET was mainly linked to monitoring of teaching within a broader evaluation system at the different governance levels of universities. But there are different perspectives about SET’s efficiency and range in the universities we visited. At GU1, teaching reports prepared by study deans are said to have changed greatly, together with the use of more sophisticated instruments for SET. However, the person in charge of SET says “SET does not yield very much and the data should be used in the faculties” (GU2.1). At GU2, the use of monitoring systems for students’ academic tracks, decisions to study certain subjects and successes or failure in classes are viewed to be a much more helpful tool for identifying problems and improving the curriculum in contrast to SET. SET is said to not contain representative data nor are students’ identities known. At GU3, SET is used to some extent in the evaluation of professors every five years, and the results enter into decision-making on additional research staff or equipment. At private GU4, SET results are used as a main source of information for annual and mid-term teaching reports to the university administration. SET is also intended to be used to give feedback to lecturers, but the possibility that it is just a ritual without any effect is also considered by interviewees.

The utility of SET for individuals mainly depends on whether they are professors with the status of civil servants or other lecturers. Professors are regulated by the Federal Civil Service Remuneration Act (1975/2002) and their salary increases almost completely depend on a salary table that does not give much weight to performance bonuses. But lecturers who are not tenured professors are more open to using SET, which can have

stronger effects on careers when used for future job applications. They view evaluations more pragmatically, as a tool for renewing their teaching contracts or for future applications. They have started to pay attention to scores and collect evaluations and then present their personal results in future hiring processes, such as at GU4 (but only marginally), GU5 and GU6. At the other universities visited, applicants have started to attach SET forms but this is not mandatory.

Statements in the newest private universities: students as clients

At all private universities, students promoted SET implementation. At GU5 and GU6, lecturers are not only made accountable to the university leadership and colleagues but students are also informed about SET results. The person in charge of SET in GU5 describes the role of students: “There is a meeting between the manager [original in English] and the entire faculty. Then all courses are looked at. There is then an average and it is watched: Which course is above average? Which course is below the average? And then, for example, the professor whose course this semester turned out so bad has to explain why” (GU5.2). When asked about the rationale for implementing SET, the person in charge explained: “One thought was always to collect the feelings, what are the expectations and needs of our students as our clients. (...). To have our clients satisfied and to share and comply with this need” (GU6.1). The reason for private universities to implement SET for all courses and distribute results to students seems to be the self-perception of the university as a rationalized organization and the role of students as clients. At public universities, not all courses are evaluated and SET results are usually not made visible to students.

Translation of SET

Some tenured German professors who cannot be dismissed or whose rank or salaries are not performance related usually constitute a group that opposes or is indifferent to SET. This group was identified at GU1, GU2, and GU3, while at GU4 they were more indifferent but not actually opposed to SET: “Some recognize it as meaningful, others as necessary, and for still others it is a necessary evil (...) they don’t see much sense in it now, they don’t attach much meaning to the judgment of students” (GU4.1). There was an attempt in the 1990s to establish SET at GU2, but “there was fierce opposition to validity. Then it was not introduced, but at some point the zeitgeist is there, and it can no longer be avoided and the opposition had to give up” (GU1.2). The person in charge of SET at GU2 explains that “German professors are rather offended, if one tries to force them. (...) If you come and say that I have to control that, you create a reactivity that wouldn’t be there by itself. They do not want to be excluded, they do not want to be controlled or manipulated (...) freedom of research and teaching is so anchored in the constitution and that is a control limit” (GU2.1). At GU3, criticism has also changed over the decades: nowadays, SET hardly produces any open conflicts about autonomy and data protection in teaching, as was the case earlier. Even if discourses on academic autonomy affect the translation of SET in German universities, the more frequent implementation and different intensity of the debates on its implementation shows in toto an increasingly institutionalized practice.

Colombia

Diffusion

The implementation date of SET in the universities we visited is between 1998 and 2008. All Colombian universities we visited, private and public, applied SET for all courses.

[Table 3 Student evaluation of teaching in Colombian universities]

Statements in public universities

The implementation of SET in the public sector is related to changes in salary rules contained in Decree 1279 (Ministerio de Educación Nacional 2002, § 14). This gave universities the option to award additional salary points for good teaching. Academic ranks depend on titles and experience, and aside from these additional points, salary increases are mostly dependent on scientific productivity as established by Decree 1279. In turn, legal compensation scales connected to performance measurement do not apply to private universities.

The relative autonomy of each public university to apply Decree 1279 has allowed them to link SET to different grades of the teaching rank and therefore, compensation. In theory, SET could be also used to dismiss professors, but in practice both processes seem to be loosely coupled either because professors are fired despite good scores or because low scores are overlooked. CU3 has twice been obliged to reemploy professors when their performance measures, including SET scores, were not used objectively in their dismissal.

Public universities may not use SET for hiring purposes. The person in charge of SET at CU1 says that when a professor asks for his or her evaluations, he responds “This does

not represent any certification or anything that you may carry as a letter of recommendation” (CU1.1). At CU2, results are published on bulletin boards. Accreditation is also viewed as historically linked to SET.

Statements in private universities

At CU5 and CU6 SET is also viewed as strongly connected to accreditation: “We started the institutional self-evaluation for accreditation purposes and in 1997 due to accreditation we needed to have some basic elements of comparability (...) and it was there that it began, to give guidelines on the quality of teaching performance” (CU6.1).

Interviewees from private CU4 did not see this linkage to accreditation but looked to universities in the US for inspiration, in particular Purdue University and the Student Instructional Report (SIR) and the SIR-II SET system available from the Educational Testing Service. Only at CU4 did students play an important role in framing the questions, in particular the “treatment [of students] was fundamental” (CU4.1). The person in charge of SET justified students’ role because: “We’re never going to do it right if we’re not going to be accountable to the student. Accountability to the student is very important” (CU4.1). At CU4, students may consult quantitative results on each course to decide whether to take it. In 2011, two of the universities visited created a group together with 11 other private universities called the “Higher Education Teaching Assessment Network”. This is a forum where they meet regularly to share their practices and evaluation forms for SET.

In the three private universities visited, low SET scores were viewed as directly coupled to the dismissal of persons and advancement of academic careers. Private CU4 establishes that a minimum teaching score should be achieved by each faculty member and that scores below this standard lead to a special assessment by peers and to dismissal if the assessment

is not “satisfactory teaching”. Each faculty can also decide to apply other data collection methods besides the survey, which is mandatory. At CU5, “students’ performance evaluation” has been mentioned since 2008, but its precise use is not regularized. Consequently, its effects on the renewal of short-term contracts for two academic periods, which is possible under Colombian labour law (Congreso de Colombia 1990) and applies to all professors at the university, are open to the subjectivity of deans and program directors.

At CU6, SET is linked to promotion and dismissal according to a qualitative scale of four levels of the Teaching Statute, which has been in effect since 2000 and was further regulated in 2017 by a Policy for Lecturers’ Development. However, according to professor CU6.2, the procedure is applied differentially to the interpretations of the highest levels of the administration led by a Rector who must be a Catholic priest, and evaluations are even not given to lecturers before dismissal. Private universities in Colombia are completely autonomous because the Colombian Constitution secures “university autonomy” (Constitución Política de Colombia 1991, § 69) at the level of the university but not academic freedom for individual academics who can be dismissed at any time.

Private universities also use SET for renewal of contracts. With adjunct lecturers, the person in charge of SET at private CU4 explains,

“The thing is sharper. The lecturers are renewed every semester, a bad grade can lead to decisions not to renew easily. (...)

Interviewer: Have you seen cases?

Yes (...) What I was saying about the professors seems to me to be very problematic.

This has to be handled with the limitations that it has, and the limitations that it has are very big” (CU4.2).

Translation of SET

In the public sector, labour unions at CU2 and CU3 have a special section for lecturers and one labour union for lecturers has occasionally criticized SET in 2011 and 2015. At CU2, the faculty union supported a claim by a professor who complained about the personalized use of SET in not renewing his contract. Faculty unions at CU3 demanded in 2018 that there be due process and effective application of the evaluation processes for the case of adjunct faculty so that they are not interpreted arbitrarily or omitted, given that SET scores were the only source of information for renewal.

The lack of regulations for contracting and dismissing these lecturers is accompanied by accusations of corruption and nepotism: “Lecturers with poor grades are not contracted again. But it all depends on personal relationships, if I’m a friend of the department director, I go on as normal and unthreatened”, explains CU2.1, the person in charge of the procedure at CU2. A professor from public CU3 explains that SET is not popular among lecturers: “It is not a peaceful issue and it is not considered that the current process is the most suitable or the most adequate to evaluate a professor” (CU3.2).

Lastly, concerns about bias and discrimination were only present in CU4, where the person in charge reformed the procedure in 2016 after internal research showed that female lecturers tended to have lower scores and female students tended to give higher scores. Class size and being in the first semester had some effect on the scores students gave, but the strongest bias was associated with the grades the students received (CU4.2). The chosen solution to alleviating bias has been technological rather than a complete review of the process: After this diagnosis, a procedure was introduced that automatically generates alternative reports that give an option to control for grades and extreme values. Also, questionnaires are now modular with questions common to the whole university. Such

technological solutions come with a cost for the university: nine persons are solely dedicated to the application of SET within the evaluation section of the university that has 40 full-time staff. In Colombia, the translation of SET also encounters methodological discourses on the validity of instruments.

Discussion

A major finding of our study is that SET has been implemented in all the universities we visited. SET is being applied in universities across national boundaries. It was initiated in the US during the 1950s by students and adopted in the 1970s by universities, while German and Colombian universities started implementing SET during the 2000s. Universities in the public *and* the private sector in countries with different higher education traditions have adopted this procedure—albeit to different degrees. Four of our German cases, three public universities and the oldest private one we visited, did not employ SET for all courses.

The spread and adoption of SET are remarkable trends and the grade of diffusion is more than we initially expected. SET has diffused into every organization we visited. We anticipated that SET may have been rejected or may have dissipated with time and that at least some of the universities we visited would not employ this procedure. But this is not the case. SET is present in universities from different traditions and sectors at least to some extent. However, we were at the same time surprised that SET is a less homogeneous practice than we thought. SET is often suggested to be an archetypical institutionalized activity in the organizational structure of contemporary universities, comparable to the presence of international offices (see Zapp and Lerch 2020). But it is translated and edited into very different discourses and that changes how it is processed.

Nevertheless, our results do support the diffusion hypothesis of neo-institutionalists to a certain extent: Universities have dealt with the increasing rationalization of society observed after the Second World War (Frank and Meyer 2020) and this trend also reaches complex activities such as university teaching. The metrification within SET demonstrates the power of rationalization discourses. Even if measuring teaching through the perception of students was first demanded by students, rationalization discourses took control and SET was later framed by other actors (academics, university administration, politicians). Technocratic control of teaching within broader processes of quality assurance in universities and the tendency to improve performance and efficiency of and within organizations are symbols of this discourse (Krücken and Meier 2006; Meyer and Ramirez 2013; Ramirez 2010). This rationalization discourse shaped the positive view on SET in the US after the mid-1970s and explains why SET was adopted nationwide in the US. Later, rationalization discourses became strong inside and outside universities in societies such as Germany and Colombia. As the cultural environment pressured for the implementation of rationalizing procedures, the standardized measurement of an old practice such as teaching acquired more legitimation two decades later in these countries.

We also found further symbols that document the influence of rationalization discourses driving isomorphism occurring during the implementation of SET and other discourses that support rationalization tendencies. Another discourse that conflates with rationalization discourses is competition. Policy talk expressing that universities and academics must engage in cycles of competition can be viewed as another element of NPM and rationalization discourse that make SET appear necessary and relevant. A further example is the economic idea of students as clients at private universities, highlighting that evaluation and accountability for teaching is rational in order to win future clients and

money. The democratic discourse of the empowerment of students that need to take part in university politics is an example for a different discourse but nevertheless, if this discourse is strong in a local setting, it supports further claims of accountability, highlighting the role of students as active and influential members of the organization. The tendency to involve students in the evaluation of lecturers has been especially strong in the US, but it is also part of a global trend to value the voices and rights of individuals over the hierarchical control of other social groups (Lerch et al. 2016). Students' interests in making teaching accountable in the 1950s explain the first appearance of SET in the US, while university administration may have taken over later for management and efficiency control reasons. Later, the empowerment of students is also a driver for the development of SET in private universities in Germany and Colombia.

Other discourses do not favour the implementation (and consequently diffusion) of SET in general. When these discourses are employed during the translation of SET, they turn its implementation into an arena of ideological conflict and establish limits for rationalization. In the data, we discovered two interesting examples for competing discourses that complicate translation and lead into editing of SET. In Germany this was the endurance of the idealized autonomy of professors as an autonomous community linked to the Humboldtian ideal. As Author 1 (2020) finds, German pedagogy based on the idea of *Bildung* is still relevant to explain translation processes. In the US, the social discourse on diversity is powerful enough to challenge the implementation of SET in the isomorphic form of standardized surveys at the end of academic periods. Two public universities had already put diversity-based arguments at the base of recent reforms but the topic was also brought up by the faculty in at least one other private university. These concerns did not arise in a vacuum but were connected to a renewed valorization of diversity after the end of

the Cold War (now challenged by exclusionary rationalisms) (Jiménez and Lerch 2019). In higher education, diversity offices operating in 59% of US universities nationwide (Kwak, Gavrilu, and Ramirez 2019) both represent structural changes brought about by these narratives and reproduce diversity discourses. Now, these pre-existing discourses were adapted by opponents of SET who also use a more technical language borrowed from statisticians. In Germany, only 20% of universities have diversity offices (Oertel 2018). We infer that cultural beliefs about the value of a diversified university explain the presence of this discourse and not necessarily the appearance of empirical investigations discovering previously unrecognized biases. Nevertheless, a methodological discourse on the validity of SET is present in all three cases and influences implementation and therefore diffusion of SET.

There are also cross-national and cross-sectoral similarities and differences in driving and impeding discourses. In all universities from all sectors, SET is a long-established practice both at public and private universities linked to rationalization and almost always supported by discourses about how lecturers and universities compete with each other. The discourse on competition, though, is less prevalent in Colombia. The politicized character of Latin American higher education (Author 1 2015) may counteract the implementation of SET conflating a competition logic. Student empowerment has often legitimated the use of SET. In Colombia and Germany, student empowerment was present in the public and private sector either in terms of participation in university politics as part of a democratic discourse or in the role as clients often linked to an economic discourse related to rationalization. The German adaptation of SET at private universities is the location where the economic discourse that epitomizes the commitment to measuring a

complex practice such as university teaching linked to the empowerment of students as economic-oriented agents is most present.

Resistance against SET based on professional discourses highlighting autonomy are only present in German public universities. These academic discourses on professorial autonomy did not stop SET's implementation. Nevertheless, they explain why the use of SET is more ritualistic in Germany compared to the other countries. Academic freedom remains strong in academic discourses and reinforces scepticism about the validity of SET results.

Actors involved in the discussion and implementation of SET also differ. In the US, academic unions supporting arguments about the perils of SET for faculty diversity play a central role. The rhetoric about exclusion or invalidity is often endorsed both by academics and academic unions. In Germany, unions are generally less involved in the academic field. The civil servant position and power of German public professors vest legitimacy to autonomy debates, but their ability to influence the agenda of SET implementation is not exercised through unions since membership is prohibited for civil servants. We also did not find any kind of organized association of untenured lecturers against SET, rather the percentage of permanent jobs is the main current topic for trade unions in universities (Gassmann 2020). In Colombia, academic unions involved in the public sector are still influenced by a Latin American tradition that has emphasized universities' role in reproducing political struggle over scientific inquiry (Author 1 2015). The arguments they support are based on the idea of validity. Validity, in this context is thought of in relation to transparency and in opposition to the corrupt clientelism that favours academic oligarchs in powerful positions in the university.

Conclusions

To conclude, our examination of discourses on SET's diffusion, translation and editing contributes to further understanding these processes that are considered highly relevant in organization studies and particularly in new institutional theory. In addition to proving the diffusion of rationalization practices with an interesting example in three different systems, it contributes to analyses of mechanisms that balance the global and the local within this trend and finally lead into what Drori, Höllerer, and Walgenbach (2014) have called 'glocalization'. We highlighted that the dialectics between the diffusion of practices, local translations and editing do not maintain an equilibrium, thus endorsing the diffusion proposition: Discourses of accountability and rationalization, of competition and of more democratic ideals to empower students are employed to favour the adoption of SET.

SET as a process that highlights the pursuit of accountability also enhances competition between academics and between institutions (for instance through teaching awards and accreditation processes). But diffusion and translation are also subject to academic discourses of professorial autonomy and methodological validity and social discourses favouring sociocultural diversity. All these discourses enter universities and exert pressure when priorities are being set for establishing policies, procedures and organizational structures. These discursive trends may be characteristic for evaluation processes such as SET, but they can also be seen as an example for the symbiosis of diffusion and translation, when ideas coming from world culture mobilize local cultural forces and lead to both homogenization and heterogeneity.

Although our main objective was not establishing which SETs are more effective than others, our analysis does provide information for university authorities when implementing SET. We offer a bird's eye view of the diffusion of SET as well as on discourses that are important for diffusing and translating SET. Our findings can also be used by academics and administrative staff responsible for SET or involved in discussions on how SET should develop at their organization, since it gives a more complete picture of concerns that are articulated in different locations but are considered in each context to improve SET. Some specific practices that we described may also serve as sources of inspiration for those in charge of evaluating teaching. For example, universities still differ in the use of dossiers and peer observations, the ban on indicators, questionnaires with modular questions or the creation of reports that already statistically compute variables identified to bias results according to local investigations. Through pondering the costs and benefits of such innovations for everyday teaching, other universities have opted to maintaining a low-cost SET system that still manages to legitimize the perceived need to measure teaching and also offers a valve for students to vent their opinions. But even in these universities, SET nowadays frequently acquires a high-stakes character for untenured lecturers. Given the increasing workload of lecturers in many universities, the consequences are that these uses of evaluation results may require a level of awareness and discussion that we did not perceive in most of our visits.

Declaration of interest

The authors declare that they have no conflict of interest.

References

- Associated Students of Stanford University. 1958. *Scratch Sheet, Dept. of Special Collections and University Archives, Stanford University Libraries*. Stanford: box 3, folder 4.
- Author 1. 2015. [details removed for peer review].
- . 2020. "[details removed for peer review]."
- Ball, Stephen. 2013. *Foucault, Power, and Education*. New York: Routledge.
- Ben-David, Joseph, and Randall Collins. 1966. "A Comparative Study of Academic Freedom and Student Politics." *Comparative Education Review* 10 (2):220-49. doi: 10.1086/445217.
- BMBF. 2018. *Gut beraten durchs Studium: Der Qualitätspakt Lehre*. Berlin: Bundesministerium für Bildung und Forschung.
- . 2019. "Projekte im Qualitätspakt Lehre." Accessed December 5. <https://www.qualitaetspakt-lehre.de>.
- Boxenbaum, Eva , and Stefan Jonsson. 2017. "Isomorphism, Diffusion and Decoupling: Concept Evolution and Theoretical Challenges." In *The SAGE Handbook of Organizational Institutionalism*, edited by Royston Greenwood, Christine Oliver, Roy Suddaby and Kerstin Sahlin, 77-101. London: SAGE.
- Brittingham, Barbara. 2009. "Accreditation in the United States: How did we get to where we are?" *New Directions for Higher Education* 2009 (145):7-27. doi: 10.1002/he.331.
- Bromley, Patricia, and John W. Meyer. 2014. "'They Are All Organizations': The Cultural Roots of Blurring Between the Nonprofit, Business, and Government Sectors." *Administration & Society* 49 (7):939-66. doi: 10.1177/0095399714548268.
- Bundesrepublik Deutschland. 1975/2002. *Bundesbesoldungsgesetz*. Bonn.
- Cohen, Louis , Lawrence Manion, and Keith Morrison. 2018. *Research Methods in Education*. London, New York: Routledge.
- Congreso de Colombia. 1990. *Ley 50 de 1990*. Bogotá.
- Constitución Política de Colombia. 1991. *Constitucion Politica de Colombia 1991*. Bogotá,: Asamblea Nacional Constituyente.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields." *American Sociological Review* 48 (2):147-60. doi: 10.2307/2095101.
- Drori, Gili S., Markus A. Höllerer, and Peter Walgenbach. 2014. "Unpacking the glocalization of organization: from term, to theory, to analysis." *European Journal of Cultural and Political Sociology* 1 (1):85-99. doi: 10.1080/23254823.2014.904205.
- Ebbinghaus, Bernhard. 2016. "When Less is More." *International Sociology* 20 (2):133-52. doi: 10.1177/0268580905052366.
- Espeland, Wendy Nelson, and Michael Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113 (1):1-40. doi: 10.1086/517897.
- Foucault, Michel. 1969/2002. *Archaeology of Knowledge*. London: Routledge.
- Frank, David John, and John W. Meyer. 2020. *The University and the Global Knowledge Society*. Princeton: Princeton University Press.

- Garcia Garduno, J. M. E. 2014. "¿Para qué sirve la evaluación de la docencia? Un estudio exploratorio de las creencias de los estudiantes universitarios." *Education Policy Analysis Archives* 22 (15). doi: 10.14507/epaa.v22n15.2014.
- Gassmann, Freya. 2020. *Befristete Beschäftigung von wissenschaftlichen Mitarbeiterinnen und Mitarbeitern an Hochschulen in Deutschland – Eine erste Evaluation der Novelle des Wissenschaftszeitvertragsgesetzes*. Frankfurt am Main: Gewerkschaft Erziehung und Wissenschaft (GEW).
- Greenwood, Royston, Mia Raynard, Farah Kodeih, Evelyn R. Micelotta, and Michael Lounsbury. 2011. "Institutional Complexity and Organizational Responses." *The Academy of Management Annals* 5 (1):317-71. doi: 10.1080/19416520.2011.590299.
- Hildebrand, Milton, Robert C. Wilson, and Evelyn R. Dienst. 1971. *Evaluating University Teaching*. Berkeley: Center for Research and Development in Higher Education in University of California, Berkeley.
- Hüther, Otto, and Georg Krücken. 2014. "The rise and fall of student fees in a federal higher education system: The case of Germany." In *Students, Markets and Social Justice: higher education fee and student support policies in Western Europe and beyond*, edited by Claire Depuy and Hubert Ertl, 85-110. Oxford: Symposium Books.
- ICEland. 2020. "Information, Controlling Entscheidung." Accessed March 10. <https://iceland.dzhw.eu/>.
- IPEDS. 2020. "Integrated Postsecondary Education Data System." Accessed February 25. <https://nces.ed.gov/ipeds/use-the-data>.
- Jiménez, Jeremy D., and Julia C. Lerch. 2019. "Waves of Diversity: Depictions of Marginalized Groups and Their Rights in Social Science Textbooks, 1900–2013." *Comparative Education Review*:000-. doi: 10.1086/702606.
- Kraatz, Matthew S, and Emily S. Block. 2008. "Organizational implications of institutional pluralism." In *The SAGE Handbook of Organizational Institutionalism*, edited by Royston Greenwood, Christine Oliver, Roy Suddaby and Kerstin Sahlin. London: SAGE.
- Krücken, Georg. 2003. "Learning the 'new, new thing': on the role of path dependency." *Higher Education* 46 (3):315–39.
- . 2019. "Multiple competitions in higher education: a conceptual approach." *Innovation*:1-19. doi: 10.1080/14479338.2019.1684652.
- Krücken, Georg, and Frank Meier. 2006. "Turning the University Into an Organizational Actor." In *Globalization and Organization: World Society and Organizational Change*, edited by Gili S. Drori, John W. Meyer and Hokyung Hwang, 241-58. New York: Oxford University Press.
- Kwak, Naejin, Gabriela Gavrila, and Francisco O. Ramirez. 2019. "Enacting diversity in American higher education: reputation and professionalization." In *Universities as Agencies*, edited by Tom Christensen, Åse Gornitzka and Francisco O. Ramirez, 209-28. New York: Universities as Agencies: Reputation and Professionalization.
- Lerch, Julia C., Patricia Bromley, Francisco O. Ramirez, and John W. Meyer. 2016. "The rise of individual agency in conceptions of society: Textbooks worldwide, 1950–2011." *International Sociology* 32 (1):38-60. doi: 10.1177/0268580916675525.
- Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *The American Political Science Review*:682-93.

- Merton, Robert K., and Patricia L. Kendall. 1946. "The Focused Interview." *American Journal of Sociology* 51 (6):541-57. doi: 10.1086/219886.
- Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World society and the nation-state." In *American Journal of Sociology*, 144-81.
- Meyer, John W., and Francisco O. Ramirez. 2013. "Universalizing the university in a world society." In *Institutionalization of World Class University in Global Competition*, edited by Jung Cheol Shin and Barbara Kehm, 257-74. Dordrecht: Springer.
- Ministerio de Educación Nacional. 2002. *Decreto 1279 de 2002*. Bogotá.
- Oertel, Simon. 2018. "The role of imprinting on the adoption of diversity management in German universities." *Public Administration* 96 (1):104-18. doi: 10.1111/padm.12384.
- Ramirez, Francisco O. 2010. "Accounting for excellence: Transforming universities into organizational actors." In *Higher Education, Policy, and the Global Competition Phenomenon*, 43-58. New York: Palgrave Macmillan.
- Ramirez, Francisco O., John W. Meyer, and Julia Lerch. 2016. "World society and the globalization of educational policy." In *The Handbook of Global Education Policy*, edited by Karen Mundy, Andy Green, Lingard. Bob and Antoni Verger, 43-63. Malden: John Wiley & Sons.
- Reinholz, Daniel L., Joel C. Corbo, Daniel J. Bernstein, and Noah D. Finkelstein. 2018. "Evaluating Scholarly Teaching: A Model and Call for an Evidence-Based Approach." In *Learning Analytics in Higher Education: Current Innovations, Future Potential, and Practical Applications*, edited by Jaime Lester, Carrie Klein, Aditya Johri and Huzefa Rangwala. New York: Routledge.
- Rindermann, Heiner. 1995. "Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen." Ruprecht-Karls-Universität Heidelberg.
- . 1997. "Generalizability of student ratings of university teaching: Can the evaluation results of one course be transferred to other courses of the same university teacher or to courses with equal content but different teachers?" *Psychologie in Erziehung und Unterricht* 44 (3):216-34.
- Sahlin-Andersson, Kerstin. 1996. "Imitating by editing success: The construction of organizational fields." In *Translating organizational change*, edited by Barbara Czarniawska and Guje Sevón, 69-92. Berlin: de Gruyter.
- Scott, Richard W. 1995. *Institutions and Organizations*. Los Angeles: Sage.
- SNIES. 2020. "Sistema Nacional de Información de Educación Superior." Accessed March 31. www.mineduacion.gov.co/snies/.
- Sproule, R. 2000. "Student evaluation of teaching: A methodological critique of conventional practices." *Education Policy Analysis Archives* 8.
- Stanford Daily. 1958. "Sibley Petition Ruled 'Out of Order'." In *Stanford Daily*, 2. Stanford.
- University of Southern California. 2019. "Teaching Evaluations Update." Accessed January 10. <https://academic senate.usc.edu/teaching-evaluations-update/>.
- Villavicencio, Judy Elena. 2008. "Teachers performance evaluation: the experience of the Dentist Academic Program of the Universidad del Valle." *Colombia Medica* 39 (1):57-63.
- Waeraas, Arild, and Jeppe Agger Nielsen. 2016. "Translation Theory 'Translated': Three Perspectives on Translation in Organizational Research." *International Journal of Management Reviews* 18 (3):236-70. doi: 10.1111/ijmr.12092.

- Wedlin, Linda, and Kerstin Sahlin. 2017. "The Imitation and translation of management ideas." In *The SAGE Handbook of Organizational Institutionalism*, edited by Royston Greenwood, Christine Oliver, Roy Suddaby and Kerstin Sahlin, 102-27. London: SAGE.
- Wellman, Barry, and Keith Hampton. 1999. "Living Networked On and Offline." *Contemporary Sociology* 28 (6). doi: 10.2307/2655535.
- Wolbring, Tobias. 2010. "Attraktivität, Geschlecht und Lehrveranstaltungsevaluation: Eine Replikationsstudie zu den Befunden von Hamermesh und Parker (2005) und Klein und Rosar (2006) mit Hilfe von Individualdaten." *Zeitschrift für Evaluation* 9 (1):29-48. doi: 10.1007/s11162-017-9452-4.
- Zapp, Mike, and Julia C. Lerch. 2020. "Imagining the World: Conceptions and Determinants of Internationalization in Higher Education Curricula Worldwide." *Sociology of Education* 93 (4):372-92. doi: 10.1177/0038040720929304.

Supplementary material

Annex 1 Categories of analysis

1. Historical context:
 - a. country
 - b. sector: public, private
 - c. religious affiliation: secular, Catholic
 - d. year of implementation
 - e. year of opposition:
 - f. implementation: mandatory,
 - g. initiators: students/administration/unions
 - h. application to some/all courses,
 - i. regulations: policy, evaluation statute, professorial rules
 - j. actors: national/regional ministry of education, courts, students, university administration, professors, assistant professors, lecturers/adjunct, teaching/didactic experts, teaching evaluation experts, economists, statisticians, faculty unions, national/international experts, companies/software, partner universities,
 - k. uses: dismissal (not tenure), monitoring, hiring, compensation/promotion, teaching award, aggregated indicators, accreditation, feedback to lecturer,
 - l. body: dean for teaching and learning, vice-president teaching and learning, center for teaching and learning, office of the register, quality assurance office, position for teaching evaluations.
2. Rationale for implementation:
 - a. application: frequency (all/some courses), online/paper, central/decentral, mid-term
 - b. access: students, colleagues
 - c. emphasis: name, constructs, type, teaching-learning, experience, client/satisfaction, customizable questionnaire, customizable reports, qualitative methods are promoted, teaching dossier, number of questions, scale
 - d. statements: rationalization, competition, student empowerment, academic freedom, invalid, diversity
 - e. high-stakes: all/some lecturers
 - f. perceived validity: class size, diversity of professors, average, perceived as invalid/ritual, class difficulty, student immaturity, course level, variation across items, response rate, extemporaneous, required for inscribing courses
 - g. other word associations.
3. Rationale for opposition:
 - a. problems: pedagogical diversity, grade inflation
 - b. critiques,
 - c. reforms