How do signals of academic performance vary across disciplines? Evidence from a survey experiment among university professors in Germany

Petzold, Knut; Netz, Nicolai

Preprint / Preprint
Zeitschriftenartikel / journal article

[Accepted for publication in Soziale Welt]

How do signals of academic performance vary across disciplines? Evidence from a survey experiment among university professors in Germany

Knut Petzold
Zittau/Görlitz University of Applied Sciences (HSZG)
Furtstraße 2, 02826 Görlitz, Germany
Knut.Petzold@hszg.de
https://orcid.org/0000-0003-1092-7718

Nicolai Netz
German Centre for Higher Education Research and Science Studies (DZHW)
Lange Laube 12, 30159 Hannover, Germany
netz@dzhw.eu
https://orcid.org/0000-0002-7272-3502

Abstract: While recent research has investigated what signals of academic performance govern academics' access to professorships, whether the power of such signals varies across disciplines has to date hardly been examined. We argue that the signaling power of academic achievements depends on the discipline-specific degree of standardization of research and on the spatio-temporal universality of research objects. Using a factorial survey experiment with Germany-based university professors of German studies, selected social sciences, and chemistry, we investigate the suitability of fictitious candidates for a tenured professorship ($N_{respondents}$ = 874, $N_{vignettes}$ = 6354). Across disciplines, we find that the formal qualification, publications and citations, and teaching experience are of primary importance for being considered suitable for a professorship, whereas international experience and connectivity are less important. Cross-level interaction analyses based on the responding professors' discipline reveal that the formal qualification is valued most in German studies and least in chemistry. For third-party funding, we find the opposite pattern. International publications and citations are similarly important in the social sciences and in chemistry, but less important in German studies. Teaching experience is rewarded equally in all disciplines. In sum, our study provides first systematic evidence for the German academic system of how the signaling power of academic achievements varies across the humanities, social, and natural sciences.

Key words: academic career success, professor, qualification, signaling, factorial survey, vignette study

Acknowledgements: We gratefully acknowledge financial support by the German Federal Ministry of Education and Research (grant number: 01PQ16002). Moreover, we thank two anonymous reviewers, Christiane Gross and Ann-Kristin Kopp for valuable feedback on our article. The Stata do-file used for our analyses is available via the DZHW Research Data Centre: https://doi.org/10.21249/DZHW:petzold2023:1.0.0.

1. Introduction

From the perspectives of the sociology of education, labor markets, and science, it is highly relevant to understand the criteria that allow academics to become professors. While recent research has made great progress in identifying criteria that govern access to professorships, it has not yet sufficiently examined how the value attached to specific signals of academic performance varies across academic disciplines.

Several studies exist that set out to identify the major criteria that influence access to professorships within single academic disciplines. The disciplines examined include political science (e.g., Habicht et al., 2021; Plümper & Schimmelfennig, 2007; Schröder et al., 2021), sociology (e.g., Lutter & Schröder, 2016), psychology (e.g., Abele-Brehm & Bühner, 2016; Lang & Neyer, 2004), economics and business administration (e.g., Schulze et al., 2008), life sciences (e.g., Jonkers, 2011), and biology (e.g., Lawson & Shibayama, 2015). Further studies cover several disciplines (e.g., Auspurg et al., 2017; Carlsson et al., 2021; Ceci, 2018; Ceci & Williams, 2015; Cruz-Castro & Sanz-Menéndez, 2010; Filandri & Pasqua, 2021; Gross et al., 2008; Jungbauer-Gans & Gross, 2013; Sanz-Menéndez et al., 2013; Weisshaar, 2017; Williams & Ceci, 2015). However, most of the latter studies concentrate on differences in the odds of attaining a professorship contingent on the academic discipline. While they occasionally touch upon disciplinary differences in the relevance of specific determinants of gaining a professorship, most notably regarding the publication record, they do not focus on such disciplinary differences. In particular, they tend not to develop and test plausible theoretical explanations for potential disciplinary differences in the signaling value of specific academic achievements.

Previous research has also not sufficiently acknowledged that career success in academia does not depend solely on the characteristics of candidates for professorships. Rather, other academics in gatekeeping positions—usually professors—evaluate candidates for professorships depending on their own background, and thereby produce discipline-specific logics and traditions in academic career success.

Moreover, we argue that societal developments have gradually changed the relevance of the criteria that make academics suitable for professorships. In addition to conventional academic achievements, such as the formal qualification, the publication record, and teaching experience, 'novel academic achievements' have gained importance, such as third-party funding (Abele-Brehm & Bühner, 2016; Lawson & Shibayama, 2015; Schröder et al., 2021) as well as international mobility, connectivity, and visibility (Geuna, 2015; Hamann & Zimmer, 2017; Netz et al., 2020).

As elaborated in the theory section, both conventional and novel academic achievements can be theorized as signals of academic performance. They should display the suitability of potential candidates for professorships in all academic disciplines. However, the signaling power of these achievements is likely to vary across disciplines. The literature on disciplinary academic cultures illustrates that disciplines differ regarding the degree of standardization of research and regarding the spatio-temporal universality of the research objects under investigation (Becher, 1994; Biglan, 1973; Simpson, 2017). In the humanities, for example, research usually focuses on specific epochs and regions, so that study designs are less standardized. Conversely, the natural sciences examine more general research objects, which concern the entire natural world. Therefore, they tend to follow highly standardized criteria for assessing scientific quality. Consequently, the signaling power of formally similar academic achievements may well differ across academic disciplines.

Previous (quantitative) studies on access to professorships also have methodological shortcomings. They typically rely on survey data that describe the careers of academics (e.g., Jungbauer-Gans & Gross, 2013; Plümper & Schimmelfennig, 2007; Schulze et al., 2008), register data of academics (e.g., Lang & Neyer, 2004), information from public websites (e.g., Habicht et al., 2021; Lutter & Schröder, 2016; Schröder et al., 2021), or processual data from appointment committees (e.g., Auspurg et al., 2017).

Data from such observational studies (Rosenbaum, 2010) often suffer from potential endogeneity bias, which results from self-selection of the examined individuals into the sample. They are also characterized by confounder problems resulting from unobserved heterogeneity between the examined individuals, implying that the possibilities of causal inferences are limited (Rubin, 2008). Some studies also apply experimental designs to investigate access to academic positions, not suffering from these problems. However, the research focus differs from ours in that the designs are used to study gender effects in particular (Carlsson et al., 2021; Ceci, 2018; Ceci & Williams, 2015; Solga et al., 2023; Williams & Ceci, 2015).

Against this background, we examine access to professorships using a factorial survey experiment administered to Germany-based university professors of German studies, selected social sciences (political science and sociology[1]), and chemistry. To each professor, we randomly presented fictitious possible candidates for professorships, thereby varying different ascriptive and meritocratic characteristics, i.e., gender, formal qualification, publications and citations, teaching experience, third-party funding, international mobility experience, and cross-border cooperation experience. We measured professors' judgments of the suitability of the presented candidates for tenured professorship at a German university. Due to our design, we can estimate both unbiased direct effects of the candidates' characteristics and cross-level interaction effects with the responding professors' discipline. Based on this design, we can compare the signaling power attributed to major academic achievements across exemplary disciplines of the humanities, social sciences, and natural sciences.

We proceed by developing theoretical thoughts on why the value of the discussed signals of academic performance should vary across the exemplary academic disciplines. Thereafter, we elaborate on our factorial survey design, the sample of responding professors, and the estimation methods. We then present our empirical results, before discussing the main contributions, limitations, and implications of our study.

2. The discipline-specific value of signals of academic performance

Academic performance is characterized by the fact that new knowledge is generated and disseminated. Although research areas, paradigms, theoretical approaches, research methods, and resources differ substantially across disciplines, research processes are usually characterized by contingency, which academics need to deal with productively. In our view, academics' skills and abilities to cope with such contingent research processes determine their academic performance. Such skills and abilities include, for example, intelligence, creativity, a systematic way of working, diligence, and resilience. Yet, academic performance is not only reflected in the research output itself, but also in its reception by other academics, and in academics' ability to build networks and collaborate with others. Professional contacts not only reflect integration into the scientific community, but may also generate important resources (e.g., Granovetter, 1973; Lin, 1999). Such resources can, in turn, increase academic performance. While a positive reception is more likely when academics are innovative and attract attention, establishing networks requires, for example, communication skills, trustworthiness, and loyalty.

When it comes to evaluating candidates' suitability for a professorship, the evaluators are interested in precisely such qualities of potential candidates. The suitability for a professorship should thus be the greater, the stronger an academic's performance is. However, such qualities are not easily directly observable, if at all.

[1] Our survey also targeted professors of geography, which we did not include in this analysis because geography comprises very different disciplinary cultures, making it hard to compare to sociology and political science, which are more homogeneous in many respects.

In this context, signaling theory addresses the fundamental communication problem of how a receiver (in our case a professor) can establish whether a sender (candidate for a professorship) is telling the truth about his or her qualities, and, relatedly, how a sender can persuade the receiver that he or she is telling the truth. To this end, a connection is established between the sender's unobservable traits and his or her observable features. Since the inception of signaling theory in labor economics, the primary quality to be displayed refers to an employer's productivity (Akerlof, 1970; Bills, 2003; Bills et al., 2017; Spence, 1973; Stiglitz, 1975). However, the behavioral and social sciences have further developed signaling theory to include features beyond an individual's productivity (e.g., Podolny, 2005; Posner, 2000; Searcy & Nowicki, 2005). From a broader point of view, any feature intentionally displayed for the purpose of convincing a receiver of a sender's desired quality can thus be considered a signal (Gambetta, 2009).

In the present study, displayed signals refer to qualities upon which the candidate's academic performance rests. It is in the interests both of candidates with the desired qualities and of professors that these qualities are truthfully displayed. From the perspective of signaling theory, the solution is that only those candidates with specific qualities will try to signal them through observable properties, provided that the signals are cheap enough for candidates possessing such qualities to acquire and emit, but too costly for those candidates without them. In this framework, academic achievements meet the essential requirements needed to function as signals of academic performance: In a perfectly separating case, all candidates with the unobservable traits will be divided from those without them by being able to emit signals of academic performance (separating equilibrium). Conversely, if both candidates with and without the qualities of interest, or none of these groups, can afford to acquire and emit the signals of academic performance, they become uninformative (pooling equilibrium). Finally, if a certain proportion of the non-quality candidates emits the signal in addition to the quality candidates, the signals do not conclusively reflect the qualities in question (semisorting equilibrium).

However, what counts as a signal and what makes it more or less costly for different types of senders depends on the specific context (Gambetta, 2009): The power of signals is not only determined by the cost of acquiring them but also by the normative systems of the senders and receivers. Successful signals are constrained by what is accepted by tradition. In this regard, it is necessary to consider the domain—and in our case the scientific discipline—in which signals are acquired, displayed, and received.

Based on the literature on disciplinary academic cultures (Becher, 1994; Biglan, 1973; Simpson, 2017), we therefore argue that the power of signals is likely to vary across academic disciplines. In our analysis, we consider German studies, sociology and political science, and chemistry as specific representatives of the humanities, the social sciences, and the natural sciences. As elaborated in the introduction, we distinguish between conventional and novel academic signals of academic performance.

2.1. Conventional signals of academic performance

2.1.1. Qualifications

Completing an academic qualification process is a well-established signal of academic performance. Importantly, disciplines differ in the degree of standardization regarding how research is conducted (Biglan, 1973; Simpson, 2017): While there is a high level of agreement on standardized criteria for assessing scientific quality in the natural sciences, the social sciences are characterized by a greater variety of epistemological paradigms. Research in the humanities is comparatively object-oriented, so that the assessment of scientific quality depends more on relevant experts in the research field, who make their evaluations with respect to research objects within formal qualification procedures.

Relatedly, the natural sciences are more internationally oriented, so that the signaling power of country-specific formal qualifications might be comparatively weak in this discipline—even if they are assessed

in the country where they were acquired. For instance, a habilitation, which is uncommon in many countries, is likely to have a much lower signaling power in the natural sciences than in other disciplines.

Following these arguments, we assume that formal qualifications should have the strongest signaling power in German studies, followed by the social sciences, and then by chemistry (hypothesis 1).

2.1.2. Publication record

Scientific publications are a core signal of academic performance (Habicht et al., 2021; Jungbauer-Gans & Gross, 2013; Long et al., 1993; Lutter & Schröder, 2016; Sanz-Menéndez et al., 2013; Schulze et al., 2008). They are an essential part of the academic production process and usually the result of a successful research process, which requires qualities related to academic performance. Therefore, a high level of publication activity should serve as a signal of academic performance in all disciplines. As the signaling value might depend on the type of publication, we differentiate between German and international publications in our empirical analysis.[2]

Research in German studies usually focuses on specific epochs and regions within the German-speaking cultural context. Research in chemistry, by contrast, is typically quite universal, so that research laboratories around the world work on similar research questions. The social sciences comprise research fields that can be defined as regional and epochal as well as universal, in that they sometimes also concern the entire humanity (Becher, 1994; Biglan, 1973; Simpson, 2017).

Accordingly, we assume that the signaling value of German publications is strongest in German studies, moderate in the social sciences, and weakest in chemistry (hypothesis 2a). In contrast, international publications should have most signaling value in chemistry, a moderate value in the social sciences, and least value in German studies (hypothesis 2b).

2.1.3. Teaching experience

In addition to research, teaching is a core task of professors at German universities. The ability to communicate theoretical approaches, methods, and findings to students and doctoral candidates and to integrate insights from current research into teaching represents a separate area of academic performance. Because teaching is essential for maintaining any discipline, we do not expect any differences in the signaling value of teaching experience between disciplines (hypothesis 3).

2.2. Novel signals of academic performance

2.2.1. Third-party funding

Scientific activities are always associated with financial costs, which require funds. Nowadays, funds are increasingly being awarded to researchers through competitive procedures. Researchers must apply for third-party funding and their proposed projects are critically assessed to ensure scientific quality. Third-party funding thus requires a high degree of academic performance. In that sense, third-party funding is another signal of academic performance (Habicht et al., 2021; Schröder et al., 2021).

Importantly, disciplines differ in terms of the degree to which high-quality research depends on costly research infrastructure and technical equipment (Becher, 1994; Biglan, 1973; Simpson, 2017): Research in the natural sciences is highly dependent on research infrastructure and technical equipment, whereas in the humanities, researchers mainly need access to their primary objects of investigation, which are increasingly available online. In the social sciences, large-scale data collection may require substantial

[2] For disciplinary comparisons, it would also be relevant to compare the relative signaling power of different numbers of publications, publication quality, and publications formats, e.g., books versus journal articles. As we did not consider these dimensions in our experimental design for practical reasons, they will have to be considered in further research.

funding, but large parts of social science research can also be carried out with small samples, or even without any empirical design, and therefore entail a comparatively low financial burden.

We therefore expect the strongest signaling effect of third-party funding in chemistry, followed by the social sciences, and the weakest effect in German studies (hypothesis 4).

2.2.2. International mobility, connectivity, and visibility

International mobility, connectivity, and visibility are also associated with academic performance (Cruz-Castro & Sanz-Menéndez, 2010; Franzoni et al., 2014; Netz et al., 2020). Stays in another country are costly in terms of the monetary, organizational, social, and psychological burdens. Yet, the returns include the acquisition of specialized knowledge and new contacts, which can promote research activities (Aman, 2020; Geuna, 2015). Therefore, experiences and characteristics related to international mobility and connectivity may also represent signals of academic performance.

As already discussed, the importance of internationality may depend on the spatio-temporal universality of research objects, and therefore vary across disciplines: The natural sciences tend to have universal research objects, while German studies tend to focus on research objects in German-speaking countries, and the social sciences are both universally and locally oriented.

We therefore assume that stays abroad (hypothesis 5a) and contact with scientists in other countries (hypothesis 5b) have the strongest signaling value in chemistry, followed by the social sciences, and the weakest signaling value in German studies.

Besides a high publication activity, international mobility and cooperation also tend to promote international visibility. Scholars who are internationally mobile and visible are likely to create new network ties, which are then likely to cite the work of the newly acquainted colleagues (Franzoni et al., 2014; Netz et al., 2020; Petersen, 2018). Citations, in turn, can be important signals of academic performance for gaining access to professorships (Baruffaldi et al., 2020; Schröder et al., 2021). Following the same reasoning as with stays and contacts abroad, we assume that citations in German and in international publications have a different signaling value across disciplines.

In detail, we hypothesize that a high number of citations in German publications are the strongest signal in German studies, followed by the social sciences and chemistry (hypothesis 6a). For a high number of citations in international publications, we expect the inverse pattern (hypothesis 6b).

3. Data and methods

Unlike most previous studies, we test our hypotheses using a factorial survey experiment (Auspurg & Hinz, 2015; Jasso, 2006; Rossi & Anderson, 1982). Following this approach, the values (levels) of experimental treatment conditions (dimensions) are systematically varied in the descriptions of hypothetical situations or persons (vignettes). In our full experimental design (vignette universe), all vignette dimensions are balanced, orthogonal, and thus not correlated amongst each other. To avoid the vignette dimensions being correlated with the respondents' own characteristics, the vignettes are randomly assigned to the respondents.

In our study, we have randomly presented fictitious candidates for professorships to professors at German universities to measure their judgments of the presented candidates' suitability for a tenured professorship at a German university. Due to the experimental design, we can estimate unbiased direct effects of the candidates' characteristics and their interaction with the responding professors' own characteristics, including their disciplinary affiliation (for details on the potentials and pitfalls of this design see Petzold & Netz, 2022).

With a few exceptions (Carlsson et al., 2021; Ceci, 2018; Ceci & Williams, 2015; Solga et al., 2023; Williams & Ceci, 2015), most previous studies on success in the German academic system used survey data on the careers of academics (e.g., Jungbauer-Gans & Gross, 2013; Plümper & Schimmelfennig, 2007; Schulze et al., 2008), information from literature data bases and handbooks of academics (e.g., Lang & Neyer, 2004), career and publication data available on public websites (e.g., Habicht et al., 2021; Lutter & Schröder, 2016; Schröder et al., 2021), or processual data from appointment committees (Auspurg et al., 2017). Other studies used qualitative research designs (Gross et al., 2008), thereby following different methodological foundations than quantitative studies. The advantages of such non-experimental data are that they provide information on real-world situations and, in the case of processual data, that they are non-reactive.

However, most studies using non-experimental data suggest that academics who eventually win a professorship differ in many unobserved characteristics from those who do not, so that the candidates' meritocratic and ascribed traits are probably confounded. For instance, there is evidence that academic achievements correlate with the size and nature of personal networks (Gross & Jungbauer-Gans, 2007; Lang & Neyer, 2004). Moreover, the characteristics of candidates and of their employing institutes can correlate due to the self-selection of candidates into specific appointment procedures for professorships (Auspurg et al., 2017). Due to problems of unobserved heterogeneity and potential endogeneity bias when using observational studies (Rosenbaum, 2010), causal inferences are typically associated with a higher degree of uncertainty (Rubin, 2008). We address this issue by using an experimental design that already minimizes unobserved heterogeneity bias during data collection (Jackson & Cox, 2013).

Former studies also neglect the fact that academic success results from an interaction of the candidates' signals of academic performance and their evaluation through other academics. Studies focusing on candidates' characteristics tend to capture the evaluations of other relevant academics only indirectly (except for Gross et al., 2008, who conducted expert interviews with academics). By contrast, a factorial survey experiment enables a direct and detailed investigation of professors' judgments of candidates' suitability for a professorship. The weights attributed to candidates' academic achievements can be estimated directly and independently from each other. Importantly, the survey experiment does not suffer from a survivor bias, as it generates data on the fictitious candidates independently of whether they are eventually considered suitable for a professorship or not. For these reasons, our experimental design produces results with a high internal validity (Mutz, 2011).

3.1. Experimental design

To avoid overly complex decision situations, factorial surveys can only consider a limited number of influencing factors. Still, the presented vignettes should contain enough information to capture the theoretically most relevant factors influencing the respondents' judgements (Auspurg & Hinz, 2015).

Based on the results of previous studies on academic career success (Abele-Brehm & Bühner, 2016; Baruffaldi et al., 2020; Cruz-Castro & Sanz-Menéndez, 2010; Jungbauer-Gans & Gross, 2013; Lang & Neyer, 2004; Lutter & Schröder, 2016; Sanz-Menéndez et al., 2013; Schulze et al., 2008; Williams & Ceci, 2015), we varied the characteristics of the fictitious candidates across ten dimensions, which comprised between two and four levels (Table 1). In detail, we varied the type of formal qualification[3], the relative number of German and international publications, and teaching experience as conventional academic achievements. In order to capture more novel academic achievements, we further considered

[3] The methodological literature suggests that dimensions varying on many levels may attract more attention-biasing responses (Verlegh et al., 2002). We must therefore take into account the possibility of such a number-of-levels-effect regarding the dimension of formal qualifications, which is the only dimension comprising four levels.

third-party funding, international mobility experience during the PhD and the postdoc period, contact with scientists abroad, and the relative number of citations in German and in international publications.[4]

Table 1 Variation of fictitious candidates' characteristics on dimensions and levels

TREATMENTS	FREQUENCIES	PERCENT
Conventional achievements		
Qualification		
None of the mentioned	1605	25.26
Junior professorship (evaluated)	1579	24.85
Habilitation (postdoctoral qualification)	1601	25.20
Non-tenured associate (W2) professorship	1569	24.69
Publications (German)		
Low number of German publications	3178	50.02
High number of German publications	3176	49.98
Publications (international)		
Low number of international publications	3178	50.02
High number of international publications	3176	49.98
Teaching experience		
Little teaching experience	3203	50.41
Much teaching experience	3151	49.59
Novel achievements		
Third-party funding		
Little third-party funding	3168	49.86
Much third-party funding	3186	50.14
International experience during the PhD		
PhD gained in Germany	3179	50.03
PhD gained abroad	3175	49.97
International experience during the postdoc		
Postdoc gained in Germany	3193	50.25
Postdoc gained abroad	3161	49.75
International networks		
Contact with few scientists abroad	3188	50.17
Contact with many scientists abroad	3166	49.83
Citations (German)		
Low number of citations in German publications	3148	49.54
High number of citations in German publications	3206	50.46
Citations (international)		
Low number of citations in international publications	3173	49.94
High number of citations in international publications	3181	50.06
$N_{vignettes}$	6354	100.00

Source: SciMo Survey of Professors (2018).

[4] We also varied the ascribed characteristic of candidates' gender. However, we did not consider candidates' gender in this study because the underlying mechanisms of group-based stereotyping and discrimination differ from the mechanism of performance-related meritocracy, which are relevant for academic performance. In order to reduce the complexity of our analyses and due to its subordinate empirical relevance, we also did not include the dimension of international mobility during studies in our analyses.

In order to help the responding professors evaluate the fictitious candidates and increase the explanatory power of our results, we fixed some relevant pieces of information in the vignette introduction. First, we asked the respondents to assess the fictitious candidates only on the basis of the information provided. Second, respondents had to evaluate the candidates' general suitability for a tenured professorship, independently of their fit with a concrete vacant position. Third, we made clear that we were interested in the suitability for a tenured professorship with an average infrastructure at a German university in the respondents' own discipline. Finally, we clarified that German publications mainly target a readership in Germany, while international publications target a readership both in Germany and abroad.

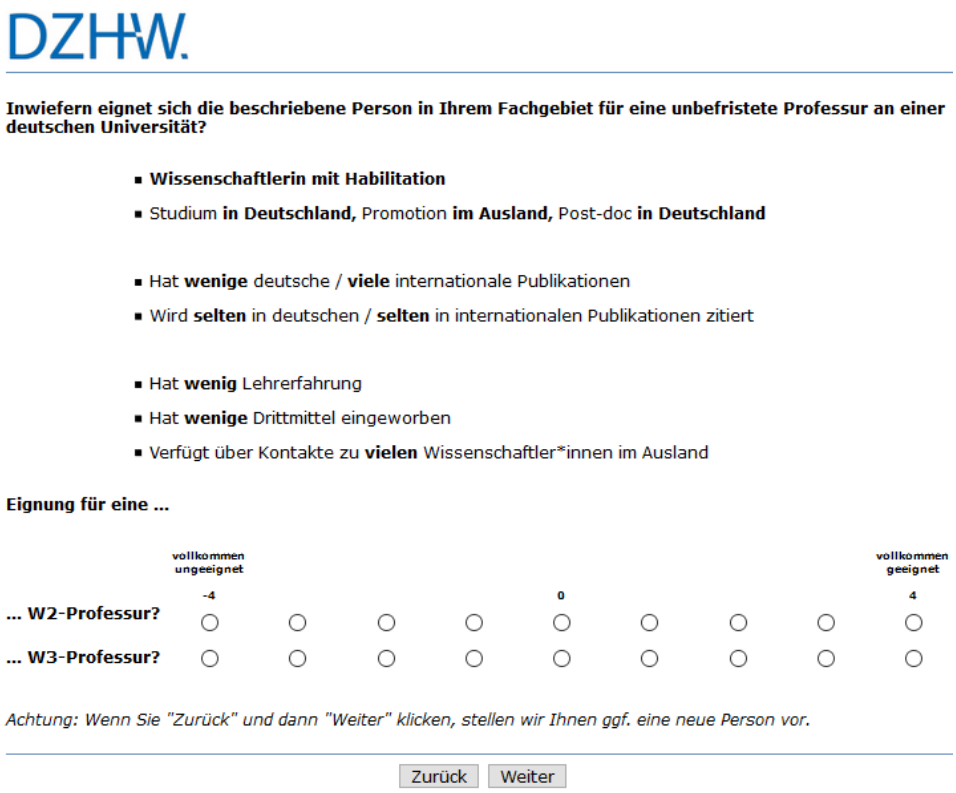
The product of the number of all levels of all dimensions (Cartesian product) reflects the maximum number of unique vignettes (vignette universe). With n = 8,192, the size of the vignette universe clearly exceeded the number of vignettes that we could present to the responding professors. Therefore, we drew a D-efficient sample of 200 vignettes (D-efficiency = 98.00). To do so, we used the modified Federov search algorithm, which sustains maximal orthogonality and level balance of all dimensions (Atzmüller & Steiner, 2010; Dülmer, 2016). In our sample, all vignette dimensions were very well balanced (Table 1) and nearly zero-correlated (Table A1 in the appendix). Also based on the algorithm, we blocked the selected vignette sample into 25 decks with eight vignettes each. Deliberate blocking allowed us to optimally balance the levels even within each deck, helping us to obtain true instead of random differences between respondents (Dülmer, 2016). Finally, we presented each respondent with a deck based on a random selection with a random order of the eight vignettes.

We implemented our factorial survey experiment using an online questionnaire (CAWI), which provided advantages over paper-based surveys regarding the random assignments and ordering of the vignettes, the recruitment of respondents (e.g., for the invitation and reminders) and convenient questionnaire completion (e.g., by enabling completion after breaks).

As Figure 1 illustrates, we asked respondents to answer the following question: "To what extent is the described person suited for a tenured professorship in your discipline at a German university?" We captured the respondents' assessment on a 9-point scale without previously specified values, as recommended in methodological literature (Sauer et al., 2011). The scale ranged from "totally unsuitable" (-4) to "totally suitable" (4). Although we captured respondents' assessment regarding both associate (W2) and full (W3) professorships, we focus on full professorships in this analysis.[5]

[5] Sensitivity analyses show that our results are very similar for both types of professorships.

Figure 1 Vignette example



Source: SciMo Survey of Professors (2018).

3.2. Sample of professors

The survey of professors was part of the project "Determinants and career effects of scientists' international mobility" (SciMo). This project was administered by the German Centre for Higher Education Research and Science Studies (DZHW) and funded by the German Federal Ministry of Education and Research (BMBF) between 2016 and 2019. The main goal of the SciMo project was to examine factors influencing scientists' international mobility and the effects of international mobility on scientists' careers (for details see Netz, 2020). However, the survey of professors was not restricted to the analysis of international mobility, but designed in such a way as to allow for broad conclusions on the relative importance of major factors influencing access to tenured professorships.

We strove for a total survey of all professors of German studies, political science, sociology, geography, and chemistry at universities in Germany. To determine this population and its composition, we used statistical information on university staff provided by the German Federal Statistical Office, which we cross-validated using data from a student information portal (studium.org). According to these sources, the population of all professors in the selected disciplines consisted of 2,729 professors in the summer semester of 2018, all of whom were invited to take part in an online survey entitled "Who is suitable for a professorship?".

Data collection took place between August and October 2018. We sent out three e-mail reminders, at one week, four weeks, and six weeks after the initial invitation. All responding professors gave their full and explicit informed consent to participate in the anonymous survey. The questionnaire was accessed by 1,162 professors (42.6 percent), and completed by 894 of them (32.8 percent). This response rate can be considered very satisfying, taking into account that highly educated individuals with highly demanding occupations are typically underrepresented in surveys and that response rates in online surveys are usually comparatively low (see also Jungbauer-Gans & Gross, 2013).

In the present analysis, we only include professors of German studies, sociology, political science, and chemistry to better account for the discipline-specific standards and to be able to investigate effects for clearly demarcated disciplines. As already explained in footnote 1, we do not include professors of geography for these reasons. Considering their many similarities, we include professors of sociology and political science in a joint category for the social sciences. Overall, 6,354 fictitious candidates were judged by 874 professors of these disciplines, including those who did not complete the entire questionnaire.[6]

Besides the address data needed for inviting the professors to our survey, we collected data from the invited professors' CVs as a supplement to the information gathered through the questionnaire. The collected CV data include information on professors' gender, year of birth, type of professorship, current university, year of obtaining the PhD, habilitation and/or evaluated junior professorship, academic discipline (disaggregated by their areas of teaching and research, the so-called *Lehr- und Forschungsbereiche*), international mobility during the studies, the PhD, and the postdoc, as well as potential awards. These data allowed us to evaluate the sample composition in comparison to the composition of the original population, and thus to assess the sensitivity of our results to processes of self-selection. Table 2 describes the target population and the estimation sample based on selected variables.

There are minor deviations between the target population and the estimation sample regarding professors' qualifications, international experience, and federal state of the current university, which are very unlikely to limit the external validity of the results. Moreover, female professors and junior professors are somewhat overrepresented. Professors from German studies and the social sciences took part more often, while professors from chemistry are underrepresented. There is thus selectivity into the estimation sample of responding professors. Yet, our experimental design itself is not biased regarding randomization and non-response. To check for the robustness of our results, we carried out a number of additional estimations (Table A3 in the appendix). First, we re-estimated our main model (Figure 3) additionally controlling for the CV characteristics of professors described in the previous paragraph, which does not lead to substantial changes in the effects of the varied academic achievements. Second, we estimated a model with fixed effects for the responding professors, which also results in almost identical estimators. Both additional models (with covariates and with fixed effects) thus indicate a successful randomization of the vignettes across the responding professors. Third, not all responding professors judged all vignettes of their deck. To check whether this non-response was systematic with regard to the content presented in the vignettes, we estimated a model with a reduced sample including only those professors for whom all eight vignette judgments were available. Again, the effects deviate only minimally and allow for the same substantial interpretations as the model with incompletely evaluated vignette decks. Therefore, we use our main model with full statistical power in the following analyses.

[6] Thirty-one respondents evaluated only one out of the eight vignettes. One hundred and forty-two respondents evaluated between two and seven vignettes (2: 23, 3: 20, 4: 14, 5: 14, 6: 14, 7: 57). Seven hundred and one respondents made evaluations of all eight vignettes of their deck (including 10 respondents we classified as satisfiers as they gave exactly the same ratings across the entire deck of vignettes). In this regard, M4 in Table A3 provides a robustness check.

Table 2 Characteristics of professors in the target population and in the sample for analyses

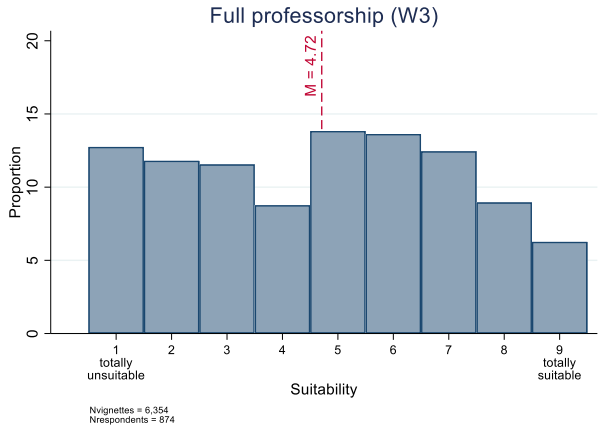
	Target Population		Sample for Analyses		
	Frequencies	Percent	Frequencies	Percent	χ^2 (p)
Gender					
Male	1905	69.81	547	62.59	31.80 (0.001)
Female	824	30.19	327	37.41	
Professorship					
Junior professor (W1)	166	6.08	78	8.92	18.17 (0.001)
(Full) professor (W2/W3)	2563	93.92	796	91.08	
Discipline					
German studies	662	24.26	267	30.55	38.31 (0.001)
Sociology	419	15.35	187	21.40	
Political Science	323	11.84	105	12.01	
Chemistry	1033	37.85	315	36.04	
Geography	292	10.70	-	-	
Qualification					
Doctoral degree	2410	88.31	795	90.96	8.75 (0.003)
Habilitation (postdoctoral qualification)	1566	57.38	478	54.69	3.81 (0.051)
Junior professorship (evaluated)	43	1.58	11	1.26	0.83 (0.361)
International experience					
As a student	722	26.46	247	28.26	2.15 (0.142)
As an academic	1726	63.25	555	63.50	0.04 (0.850)
Academic award	797	29.20	260	29.75	0.18 (0.668)
County					
Baden-Württemberg	388	14.22	131	14.99	21.23 (0.130)
Bayern	397	14.55	115	13.16	
Berlin	202	7.40	64	7.32	
Brandenburg	51	1.87	21	2.40	
Bremen	49	1.80	19	2.17	
Hamburg	81	2.97	24	2.75	
Hessen	257	9.42	73	8.35	
Mecklenburg-Vorpommern	52	1.91	14	1.60	
Niedersachsen	226	8.28	74	8.47	
Nordrhein-Westfalen	608	22.28	186	21.28	
Rheinland-Pfalz	129	4.73	49	5.61	
Saarland	26	0.95	13	1.49	
Sachsen	120	4.40	35	4.00	
Sachsen-Anhalt	50	1.83	22	2.52	
Schleswig-Holstein	22	0.81	11	1.26	
Thüringen	71	2.60	23	2.63	
$N_{respondents}$	2729	100.00	874	100.00	
$N_{vignettes}$	21832		6354		

Source: SciMo Survey of Professors (2018).

3.3. Estimation methods

Figure 2 shows the distribution of the judgments of suitability for a full professorship at a German university on our 9-point scale across all presented vignettes. There is a reasonable variance, without an overly frequent rating of the ends of the scale, indicating that the varied candidates' characteristics were relevant for the responding professors. Moreover, there is no evidence of biasing censoring effects. The rather symmetrical distribution (M = 4.72; SD = 2.45; Skewness = 0.02) further justifies the use of linear estimation models.[7]

Figure 2 Distribution of judged suitability for a full professorship across all candidates



Source: SciMo Survey of Professors (2018).

The assessment of up to eight fictitious candidates by each professor resulted in a hierarchical data structure (Hox et al., 1991; Jasso, 2006). We take this data structure into consideration by estimating random intercept fixed slope models, which account for the variation in the outcome variable between respondents (Snijders & Bosker, 2012). Because of the small size of the decks, we only estimate the intercepts with a random component. We corrected for potential remaining non-modeled heteroscedasticity through robust Huber-White standard errors (White, 1980). We estimate the following equation:

$$I_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 Z_j + v_j + \varepsilon_{ij} \qquad ; i = 1, \ldots, n; j = 1, \ldots, m$$

I_{ij}: Judgment of fictitious candidate i by responding professor j

X_{ij}: Vector of fictitious candidates' characteristics varied in vignettes

Z_j: Vector of responding professors' characteristics

v_j: Error term at responding professors' level

ε_{ij}: Error term at fictitious candidates' level

[7] Although the outcome variable differs from a normal distribution, the underlying assumption that the model residuals are normally distributed is fulfilled. We verified this through graphical analyses and a normality test (skewness and kurtosis test: adjusted overall $Chi^2 = 0.67$, $p = 0.714$). Hence, the p-values of our significance tests are likely to be valid.

We are particularly interested in how the responding professors' academic discipline moderates the weights attributed to candidates' signals of academic performance. For this purpose, we additionally include cross-level interaction terms between the fictitious candidates' meritocratic dimensions and the responding professors' discipline ($X_{ij}Z_j$). This strategy reflects a subgroup analysis with efficient estimations of vignette evaluations differentiated by all professors who belong to the same discipline.

$$I_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 Z_j + \beta_3 X_{ij} Z_j + v_j + \varepsilon_{ij} \qquad ; i = 1, \ldots, n; j = 1, \ldots, m$$

4. Empirical results

4.1. General signaling value of academic achievements

As argued above, academic achievements can be seen as signals of academic performance and thus unfold positive effects on the evaluation of the suitability for a professorship. Figure 3 shows a test of this assumption based on a joint estimation model for all covered disciplines.[8] As expected, the considered academic achievements increase the suitability for a professorship if compared to the respective reference categories. However, it becomes clear that professors consider conventional academic achievements to be more important than novel academic achievements.

A habilitation and a non-tenured associate (W2) professorship show the strongest effects.[9] These achievements are associated with an average increase of 1.4 points on the suitability scale if compared to not having any of the presented formal qualifications. A junior professorship, on the other hand, increases the suitability slightly less than a habilitation and a non-tenured associate (W2) professorship, that is, by about one scale point. A junior professorship thus has about the same effect as much teaching experience (compared to little) or as a high number of international publications (compared to a low number). Having a high number of German publications is far less effective than having a high number of international publications.

Among the novel academic achievements, much third-party funding (compared to little) and a high number of citations in international publications (compared to a low number) prove to be the strongest predictors of suitability for a professorship, with effects of around one scale point. With effect sizes of just under half a scale point, contact with many scientists abroad (compared to few) are about as important as a high number of citations in German publications (compared to a low number). The effect of a PhD gained abroad, on the other hand, is near zero and insignificant. Contrary to expectations, a postdoc gained abroad even has a significantly negative effect.

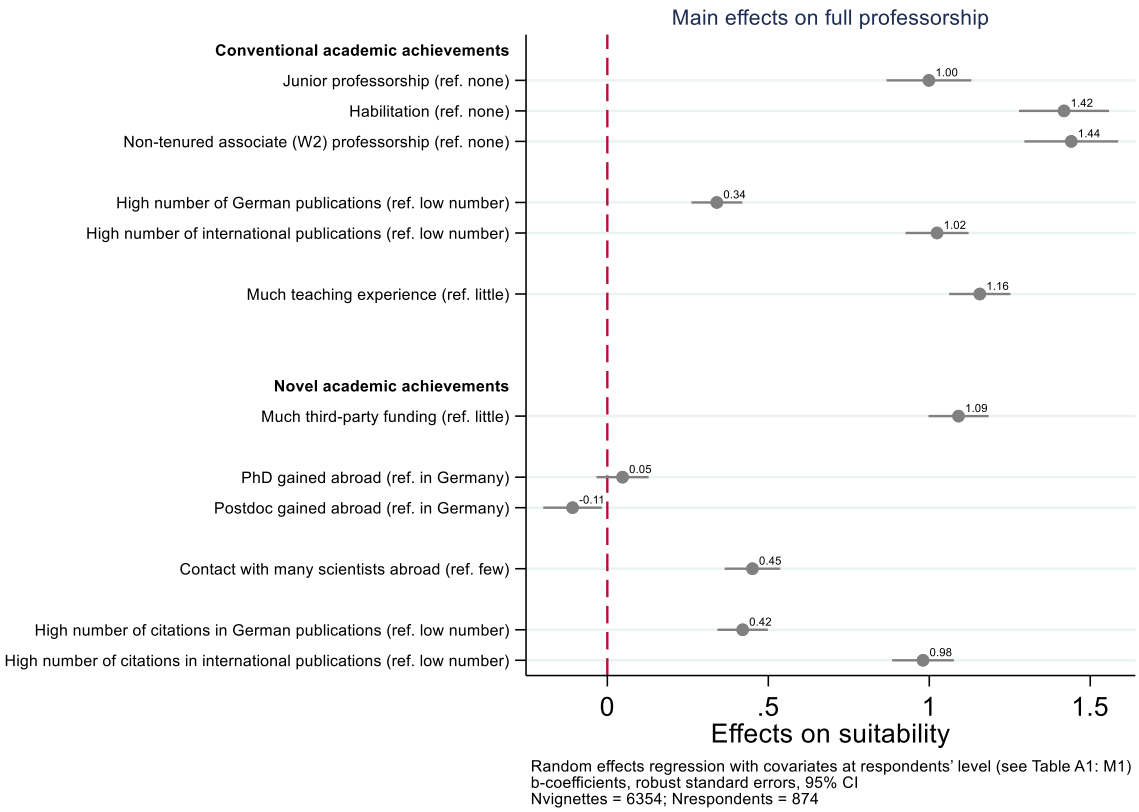
4.2. Signaling value of conventional academic achievements across disciplines

Cross-level interaction analyses reveal remarkable heterogeneity of the effects of the candidates' academic achievements across the disciplines of the responding professors. Figure 4 shows the conditional effects for the conventional academic achievements, and Figure 5 for the novel academic achievements.

[8] For the sake of an easy interpretation, we present the main results using graphical plots. Table A2 in the appendix provides detailed estimates and model information.

[9] The reference category in this dimension reflects a very low level of formal qualification (none of the mentioned qualifications). Accordingly, the importance of the effects of being habilitated and of holding a W2 professorship should not be overstressed. Instead, differences between the three formal qualifications are more informative.

Figure 3 Main effects of the varied academic achievements on the suitability for a professorship

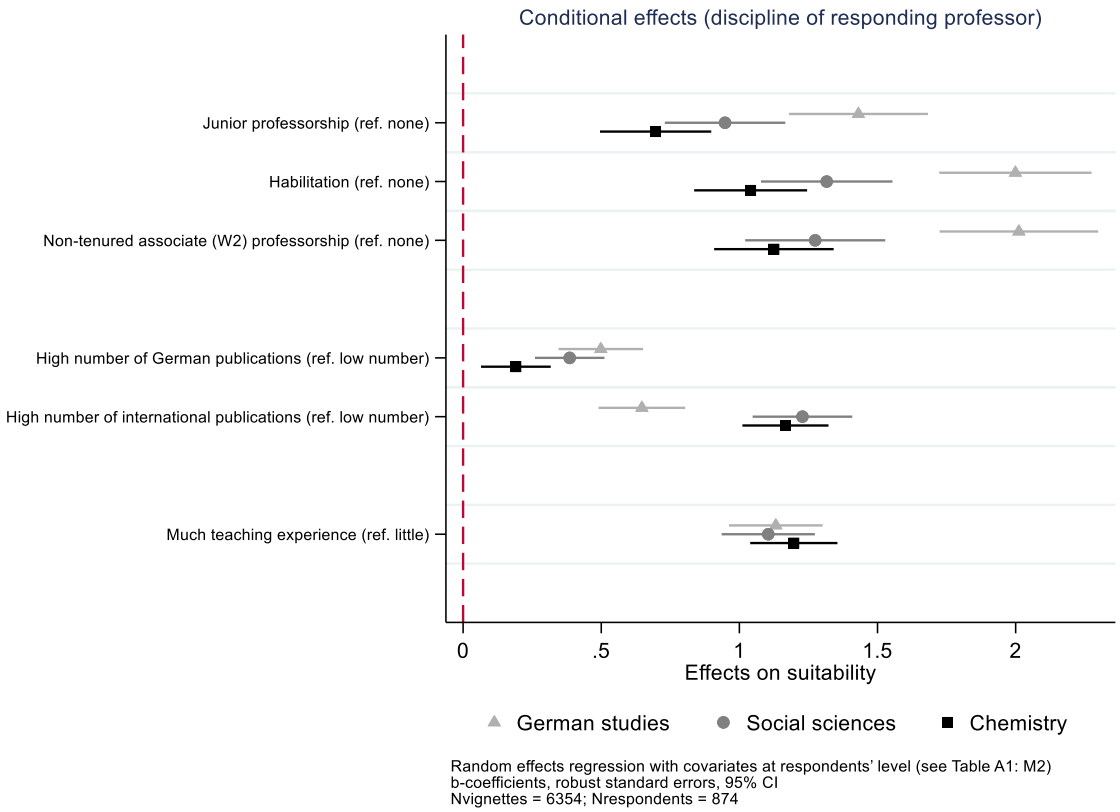


Source: SciMo Survey of Professors (2018).

We assumed that formal qualifications should have most signaling value in German studies, followed by that in the social sciences and in chemistry, the reason being differences in the criteria for assessing scientific quality in these fields (hypothesis 1). Indeed, all varied formal qualifications are more important for the judged suitability for a professorship in German studies than in the social sciences and chemistry. With two scale points, the effects of a habilitation and a non-tenured associate (W2) professorship are about twice as large in German studies as in chemistry. A junior professorship has the weakest effect in all three disciplines compared to the other types of formal qualification. However, a junior professorship still has more signaling value in German studies than a habilitation and a non-tenured associate (W2) professorship in the social sciences or in chemistry. This analysis thus clearly supports hypothesis 1.

As scientific publications are a core signal of academic performance, we expected a high publication activity to be rewarded in all disciplines. However, in accordance with the more or less universal character of discipline-specific research subjects concerning region and epoch, international and German publications may have different signaling values. Accordingly, we assumed that German publications might be stronger signals in German studies than in the social sciences and in chemistry (hypothesis 2a). As Figure 4 shows, the effect of a high number of German publications (reference: low number of German publications) is strongest in German studies, weaker in the social sciences, and weakest in chemistry. However, the difference is only significant between German studies and chemistry. Still, the relations of all effect sizes correspond to hypothesis 2a (see M2 in Table A2 for detailed effect differences).

Figure 4 Effects of conventional academic achievements conditional on academic disciplines



Source: SciMo Survey of Professors (2018).

Moreover, we expected less signaling power of international publications in German studies compared to the social sciences and especially compared to chemistry (hypothesis 2b). In fact, international publications are of much less importance in German studies than in the other two disciplines. Yet, there is no difference between the effects of a high number of international publications in the social sciences and chemistry. The signaling value of German and international publications differs only slightly in German studies, while international publications weigh more than twice as much as German publications in the social sciences and in chemistry. Regarding the comparison of chemistry and German studies, our results thus align with hypothesis 2b.

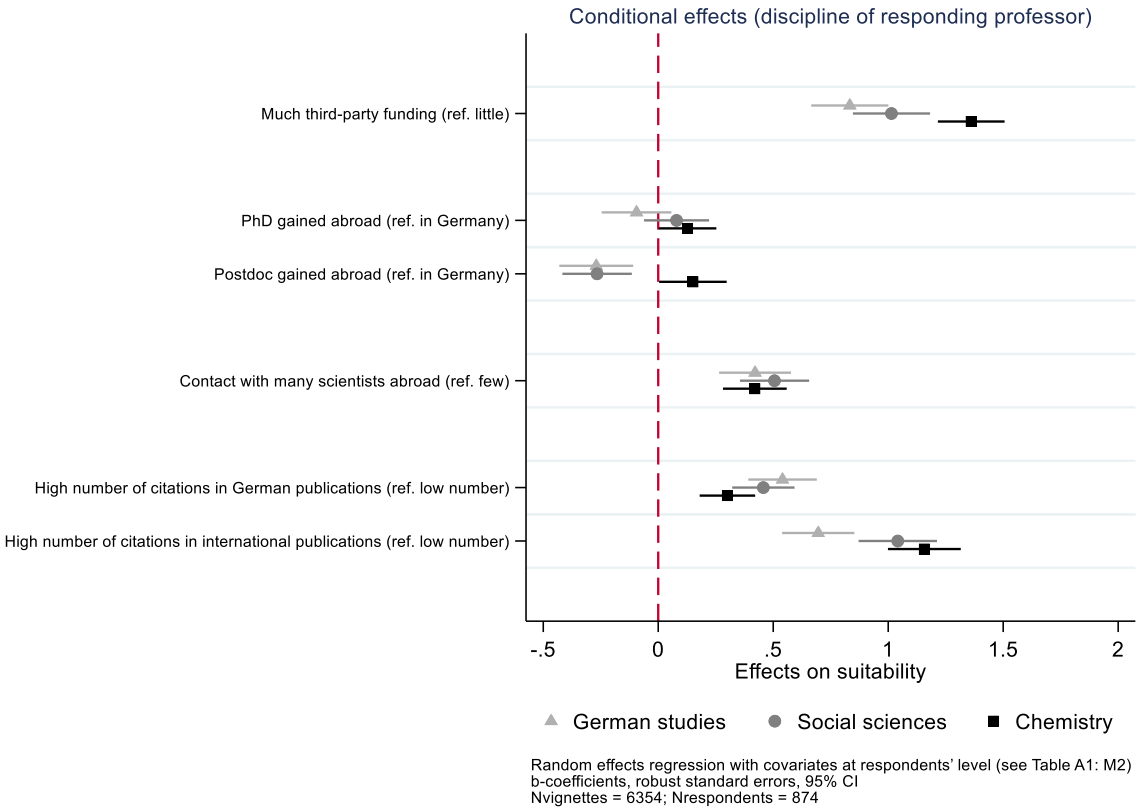
Because of its fundamental character in all academic disciplines, we did not expect any differences in the signaling power of teaching experience (hypothesis 3). The empirical analysis confirms this hypothesis. In comparison to little teaching experience, much teaching experience has a notable effect on the suitability for a professorship (more than one scale point) regardless of the evaluating professors' discipline.

4.3. Signaling value of novel academic achievements across disciplines

Third-party funding is an important signal among the novel academic achievements (Figure 5). While research in chemistry is almost impossible without generous funding of technical equipment, research in German studies primarily requires the funding of personnel and access to literature. In the social sciences, researchers may incur different amounts of cost depending on the chosen research design. We therefore expected the strongest signaling effect of third-party funding in chemistry, followed by the social sciences and German studies (hypothesis 4). In line with our expectations, third-party funding has the strongest signaling value in chemistry and the weakest in German studies; this difference is

statistically significant. Although third-party funding is considered more valuable in the social sciences than in German studies, this difference is not statistically significant. It is worth mentioning that much third-party funding shows the strongest effect of all varied achievements in chemistry. In summary, our results support hypothesis 4.

Figure 5 Effects of novel academic achievements conditional on academic disciplines



Source: SciMo Survey of Professors (2018).

Furthermore, we expected characteristics related to international mobility, connectivity, visibility and reception to serve as signals of academic performance. Once again referring to the degree of spatio-temporal universality of discipline-specific research subjects, we expected differences in their signaling values across academic disciplines.

We assumed stays abroad (hypothesis 5a) and many international contacts (hypothesis 5b) to be most important in chemistry, to be of moderate importance in the social sciences, and to be least important in German studies. However, our analyses reveal that gaining either a PhD or a postdoc in another country have (weakly positive) significant effects on the suitability for a professorship only in chemistry. In German studies and in the social sciences, a PhD gained abroad has very small and insignificant effects. A postdoc gained abroad even has a significantly negative effect in the latter two disciplines. While the effects of stays abroad in chemistry thus correspond to hypothesis 5a, this cannot be confirmed for either German studies or the social sciences.

Neither do our analyses provide empirical support for the assumed discipline-specific differences regarding the effect of contact with other scientists internationally (hypothesis 5b). The importance of

contact with many scientists abroad is weighted positively in all three disciplines, but we do not observe significant differences between disciplines in this respect.

Finally, we hypothesized that a high number of citations in German publications are the strongest signal of suitability for a professorship in German studies, followed by the social sciences and chemistry (hypothesis 6a). For a high number of citations in international publications, we assumed the inverse pattern (hypothesis 6b). In line with our expectations, the results show a pattern quite similar to the signaling value attributed to German and international publications. Citations in German publications are rewarded most by professors of German studies and rewarded least by professors of chemistry. However, these differences are only marginally significant at the five percent level. The result is clearer for citations in international publications: In German studies, a high number of citations in international publications have a similar signaling value as a high number of citations in German publications do. In the social sciences and in chemistry, a high number of citations in international publications have substantially more signaling power than a high number of citations in German publications. Overall, our findings therefore align with hypotheses 6a and 6b.

5. Discussion and conclusion

We examined how the value attributed to specific signals of academic performance varies across academic disciplines when considering the suitability of academics for a professorship. Our contribution is twofold: First, building on signaling theory, we proposed an approach that takes into account that success in academia largely depends on assessment by other academics. From this perspective, academic achievements are screened by professors in terms of their power to signal qualities that candidates' academic performance rests upon. In this context, we distinguished between conventional and novel academic achievements, assuming that they both promote the suitability of potential candidates for professorships in all academic disciplines. Following the literature on disciplinary academic cultures, we additionally argued that the signaling power of academic achievements should vary across disciplines because of discipline-specific degrees of standardization of research and the spatio-temporal universality of research objects.

Second, in contrast to most previous studies, we examined access to professorships using a factorial survey experiment, which was administered to Germany-based university professors. We randomly presented fictitious possible candidates for professorships, thereby varying major academic achievements, and measured the responding professors' judgments of the suitability of the presented candidates for tenured professorship at a German university. This research design allowed us to estimate unbiased effects of the candidates' academic achievements conditional to the responding professors' discipline. We compared the estimated effects across professors of German studies, selected social sciences (political science and sociology), and chemistry. In summary, our analyses revealed remarkable heterogeneity in the effects of the examined academic achievements across the covered disciplines.

As expected, formal qualifications do not play a prominent role in chemistry. Here, the qualification is no more important than a high number of international publications and much teaching experience. Much third-party funding is the most important criterion in chemistry, followed by a high number of citations in international publications. German publications and citations are of less importance.

In German studies, in contrast, formal qualifications, such as a habilitation or non-tenured professorship, are the most important criterion for the assessed suitability for a professorship. Publications and citations have less weight, regardless of whether they appear in German or international publication media. The relative importance of third-party funding and teaching experience is also evident in German studies.

In the social sciences, the attributed signaling values tend to range between those of chemistry and those of German studies. In terms of formal qualifications as well as German and international publications

and citations, the effects are similar to those in chemistry. Conversely, third-party funding and international mobility are evaluated in a similar way as in German studies.

There are no significant disciplinary differences regarding teaching experience and contact to scientists abroad. If the effect of international mobility is estimated net of the effects of all other dimensions, which is the case in our study by design, it hardly plays a role in the considered disciplines or is even slightly detrimental (German studies).

The facts that existing studies seldom strove for systematic disciplinary comparisons regarding the signaling value of specific academic achievements, and that many studies focused on different disciplines than the ones we examined impede robust comparisons of our results to existing ones. Broadly speaking, however, our results align with existing evidence for German academia in that the formal qualification, most notably a habilitation, has a larger signaling value in disciplines that are geared towards German society, such as German studies or law, while—especially internationally visible—publications are more relevant in the social and the natural sciences (Gross et al., 2008).[10] We also confirm research for the German social sciences that the qualification exerts a positive signaling effect even net of the publication performance (Lutter & Schröder, 2016; Schröder et al., 2021).

Moreover, our results correspond to previous evidence in that they did not reveal notable disciplinary differences concerning the signaling value of teaching experience in German academia (Gross et al., 2008).

Regarding third-party funding, our analyses tend to support recent empirical evidence for the social sciences suggesting that this is of high importance (Abele-Brehm & Bühner, 2016; Schröder et al., 2021; Solga et al., 2023), rather than older empirical evidence suggesting that it is only marginally important for obtaining a professorship (Plümper & Schimmelfennig, 2007; Schulze et al., 2008). On a broader note and beyond the comparison of disciplines, this supports our initial argument that novel signals of academic performance may have become more important over the past decades.

Finally, our results confirm previous evidence that ascriptions of internationality play a greater role in the natural sciences than in the social sciences and especially than in the humanities (Hamann & Zimmer, 2017).

Our study has several limitations, which represent starting points for future research. As with every experiment, we had to select specific theoretically relevant treatments that influence the suitability for a professorship. The existing literature has shown the chosen dimensions to be important determinants of academic career success. The theoretical proposition to understand academic achievements as signals of academic performance readily allows for extensions. Signaling theory can easily be applied to different regional and temporal contexts when it comes to displaying academic performance through observable features. Accordingly, future studies may set other priorities when varying dimensions and levels.

We studied the relative proportion of German and international publications and citations. In the future, it would be interesting to also explicitly differentiate types of publications, for example written books, editorships, and peer-reviewed journal articles. Moreover, the share of co-authorships typically differs between disciplines due to discipline-specific cooperation norms (Gross et al., 2017). It is therefore possible that the diverging importance of the number of publications corresponds to differing proportions of co-authorships across disciplines. In this respect, further research is needed considering

[10] Our results also substantiate the picture that internationally visible publications are nowadays equally important in the social sciences as in the natural and technical sciences (Abele-Brehm & Bühner, 2016; Jonkers, 2011; Jungbauer-Gans & Gross, 2013; Lang & Neyer, 2004; Lutter & Schröder, 2016; Schröder et al., 2021; Schulze et al., 2008).

the competing signaling values of co-authorships reflecting scientific cooperation on the one hand and single authorships reflecting scientific contributions attributable to individual researchers on the other. Finally, single publications can be particularly influential and generate large numbers of citations, for example when new theoretical approaches or empirical methods are successfully introduced. The role of such outstanding publications and the associated citations could also be examined in future research, including the possibility that publications presenting entirely novel approaches might—at least initially—be more difficult to publish and face a citation penalty (Wang et al., 2017).

A more nuanced analysis of different types of international mobility would also be beneficial. For instance, it is plausible that the value of academic stays abroad—and corresponding variation across disciplines—differs depending on the host country and institution. In German studies, stays in German-speaking countries may be particularly beneficial, while stays in Anglophone countries could be more relevant in chemistry, where English is the ultimate lingua franca; as the social sciences often have a regional focus, stays in countries related to the specific objects of study arguably matter. Similar patterns might be observable regarding institutional prestige. Ultimately, stays in specific countries and at specific institutions may thus be understood as a matter of (mis)fit of academic cultures. From this perspective, our finding that professors in German studies and the social sciences assess a postdoc gained abroad negatively may reflect a suspicion that candidates are insufficiently socialized in the German academic system if they completed their postdoc gained abroad.

Typically, candidates for professorships differ regarding their individual academic achievements. Someone may have published a lot and gained extensive international experience, but only have a little teaching experience and possibly no qualifications beyond the doctorate. This raises the question of whether specific signals of academic performance can be substituted by each other. For example, can the importance of international experience be substituted by a comparatively large number of international publications, or vice versa? Such substitution processes—and possible variation of these processes across disciplines—clearly deserve more attention.

We focused on academic achievements, which are usually gained through one's own efforts and thus follow a meritocratic principle. However, previous research has shown that academic careers are also determined by ascribed characteristics such as gender and immigrant background (e.g., Gross & Jungbauer-Gans, 2007; Lutter & Schröder, 2016; Solga et al., 2023; Williams & Ceci, 2015). In addition, the signaling effect of specific universities is becoming more and more differentiated. The importance of ascribed and institutional characteristics, and of potential interactions with more meritocratic academic achievements, should also be examined more closely in the future, inter alia by applying experimental research designs.

Furthermore, we only compared the signaling power of academic achievements across German studies, sociology and political science, and chemistry, which served as representatives of the humanities, the social sciences, and the natural sciences. Moreover, we focused on German academia. Due to possible discipline-specific and country-specific idiosyncrasies, the generalizability or our results is thus limited. Comparative studies including further disciplines and countries are therefore desirable.

Addressing these and other aspects may help to further understand the varying importance of academic achievements across different disciplinary, institutional, and country contexts. The present experimental study on the signaling value of specific academic achievements in German studies, selected social sciences, and chemistry has laid the foundations for this line of research.

References

Abele-Brehm, A., & Bühner, M. (2016). Wer soll die Professur bekommen? Eine Untersuchung zur Bewertung von Auswahlkriterien in Berufungsverfahren der Psychologie. *Psychologische Rundschau*, *67*(4), 250–261. https://doi.org/10.1026/0033-3042/a000335

Akerlof, G. (1970). The Market for „Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500. https://doi.org/10.2307/1879431

Aman, V. (2020). Transfer of knowledge through international scientific mobility: Introduction of a network-based bibliometric approach to study different knowledge types. *Quantitative Science Studies*, *1*(2), 565–581. https://doi.org/10.1162/qss_a_00028

Atzmüller, C., & Steiner, P. (2010). Experimental vignette studies in survey research. *Methodology*, *6*(3), 128–138. https://doi.org/10.1027/1614-2241/a000014

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. SAGE Publications.

Auspurg, K., Hinz, T., & Schneck, A. (2017). Berufungsverfahren als Turniere: Berufungschancen von Wissenschaftlerinnen und Wissenschaftlern. *Zeitschrift für Soziologie*, *46*(4), 283–302. https://doi.org/10.1515/zfsoz-2017-1016

Baruffaldi, S., Marino, M., & Visentin, F. (2020). Money to move: The effect on researchers of an international mobility grant. *Research Policy*, *49*(8), 104077. https://doi.org/10.1016/j.respol.2020.104077

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, *19*(2), 151–161. https://doi.org/10.1080/03075079412331382007

Biglan, A. (1973). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, *57*(3), 204–213. https://doi.org/10.1037/h0034699

Bills, D. (2003). Credentials, signals, and screens: Explaining the relationship between schooling and job assignment. *Review of Educational Research*, *73*(4), 441–469. https://doi.org/10.3102/00346543073004441

Bills, D., Di Stasio, V., & Gërxhani, K. (2017). The Demand Side of Hiring: Employers in the Labor Market. *Annual Review of Sociology*, *43*(1), 291–310. https://doi.org/10.1146/annurev-soc-081715-074255

Carlsson, M., Finseraas, H., Midtbøen, A., & Rafnsdóttir, G. (2021). Gender Bias in Academic Recruitment? Evidence from a Survey Experiment in the Nordic Region. *European Sociological Review*, *37*(3), 399–410. https://doi.org/10.1093/esr/jcaa050

Ceci, S. (2018). Women in Academic Science: Experimental Findings From Hiring Studies. *Educational Psychologist*, *53*(1), 22–41. https://doi.org/10.1080/00461520.2017.1396462

Ceci, S., & Williams, W. (2015). Women have substantial advantage in STEM faculty hiring, except when competing against more-accomplished men. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01532

Cruz-Castro, L., & Sanz-Menéndez, L. (2010). Mobility versus job stability: Assessing tenure and productivity outcomes. *Research Policy*, *39*(1), 27–38. https://doi.org/10.1016/j.respol.2009.11.008

Dülmer, H. (2016). The Factorial Survey: Design Selection and its Impact on Reliability and Internal Validity. *Sociological Methods & Research*, *45*(2), 304–347. https://doi.org/10.1177/0049124115582269

Filandri, M., & Pasqua, S. (2021). 'Being good isn't good enough': Gender discrimination in Italian academia. *Studies in Higher Education*, *46*(8), 1533–1551. https://doi.org/10.1080/03075079.2019.1693990

Franzoni, C., Scellato, G., & Stephan, P. (2014). The mover's advantage: The superior performance of migrant scientists. *Economics Letters*, *122*(1), 89–93. https://doi.org/10.1016/j.econlet.2013.10.040

Gambetta, D. (2009). Signaling. In P. Hedström & P. Bearman (eds.), *The Oxford Handbook of Analytical Sociology* (pp. 168-194). Oxford University Press.

Geuna, A. (ed.). (2015). *Global Mobility of Research Scientists: The Economics of Who Goes Where and Why*. Academic Press. https://doi.org/10.1016/C2014-0-00191-8

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360-1380. https://doi.org/10.1086/225469

Gross, C., & Jungbauer-Gans, M. (2007). Erfolg durch Leistung? Ein Forschungsüberblick zum Thema Wissenschaftskarrieren. *Soziale Welt*, *58*(4), 453–471. https://doi.org/10.5771/0038-6073-2007-4-453

Gross, C., Jungbauer-Gans, M., & Kriwy, P. (2008). Die Bedeutung meritokratischer und sozialer Kriterien für wissenschaftliche Karrieren–Ergebnisse von Expertengesprächen in ausgewählten Disziplinen. *Beiträge Zur Hochschulforschung*, *30*(4), 8–32.

Gross, C., Jungbauer-Gans, M., & Nisic, N. (2017). Cooperation and Career Chances in Science. In B. Jann & W. Przepiorka (eds.), *Social dilemmas, institutions and the evolution of cooperation* (pp. 165–187). De Gruyter Oldenbourg. https://doi.org/10.1515/9783110472974-009

Habicht, I., Lutter, M., & Schröder, M. (2021). How human capital, universities of excellence, third party funding, mobility and gender explain productivity in German political science. *Scientometrics*, *126*(12), 9649–9675. https://doi.org/10.1007/s11192-021-04175-8

Hamann, J., & Zimmer, L. (2017). The internationality imperative in academia. The ascent of internationality as an academic virtue. *Higher Education Research & Development*, *36*(7), 1418–1432. https://doi.org/10.1080/07294360.2017.1325849

Hox, J., Kreft, I., & Hermkens, P. (1991). The analysis of factorial surveys. *Sociological Methods & Research*, *19*(4), 493–510. https://doi.org/10.1177/0049124191019004003

Jackson, M., & Cox, D. (2013). The Principles of Experimental Design and Their Application in Sociology. *Annual Review of Sociology*, *39*(1), 27–49. https://doi.org/10.1146/annurev-soc-071811-145443

Jasso, G. (2006). Factorial Survey Methods for Studying Beliefs and Judgments. *Sociological Methods & Research*, *34*(3), 334–423. https://doi.org/10.1177/0049124105283121

Jonkers, K. (2011). Mobility, productivity, gender and career development of Argentinean life scientists. *Research Evaluation*, *20*(5), 411–421. https://doi.org/10.3152/095820211X13176484436177

Jungbauer-Gans, M., & Gross, C. (2013). Determinants of Success in University Careers: Findings from the German Academic Labor Market. *Zeitschrift für Soziologie*, *42*(1), 74–92. https://doi.org/10.1515/zfsoz-2013-0106

Lang, F., & Neyer, F. (2004). Kooperationsnetzwerke und Karrieren an deutschen Hochschulen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *56*(3), 520–538. https://doi.org/10.1007/s11577-004-0076-2

Lawson, C., & Shibayama, S. (2015). International research visits and careers: An analysis of bioscience academics in Japan. *Science and Public Policy*, *42*(5), 690–710. https://doi.org/10.1093/scipol/scu084

Lin, N. (1999). Social Networks and Status Attainment. *Annual Review of Sociology*, *25*, 467-487. https://doi.org/10.1146/annurev.soc.25.1.467

Long, S., Allison, P., & McGinnis, R. (1993). Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity. *American Sociological Review*, *58*(5), 703–722. https://doi.org/10.2307/2096282

Lutter, M., & Schröder, M. (2016). Who becomes a tenured professor, and why? Panel data evidence from German sociology, 1980–2013. *Research Policy*, *45*(5), 999–1013. https://doi.org/10.1016/j.respol.2016.01.019

Mutz, D. (2011). *Population-Based Survey Experiments*. Princeton University Press.

Netz, N. (2020). *Determinanten und berufliche Effekte der internationalen Mobilität von Wissenschaftler*innen. Schlussbericht zum SciMo-Projekt*. DZHW.

Netz, N., Hampel, S., & Aman, V. (2020). What effects does international mobility have on scientists' careers? A systematic review. *Research Evaluation*, *29*(3), 327–351. https://doi.org/10.1093/reseval/rvaa007

Petersen, A. (2018). Multiscale impact of researcher mobility. *Journal of The Royal Society Interface*, *15*(146), 20180580. https://doi.org/10.1098/rsif.2018.0580

Petzold, K., & Netz, N. (2022). Vignettenexperimente in der Hochschul- und Wissenschaftsforschung: Konstruktion, Potenziale und Fallstricke illustriert am Beispiel einer Professor*innenbefragung. In G. Brandt & S. de Vogel (eds.), *Survey-Methoden in der Hochschulforschung* (pp. 153–199). Springer. https://doi.org/10.1007/978-3-658-36921-7_7

Plümper, T., & Schimmelfennig, F. (2007). Wer wird Prof – und wann? Berufungsdeterminanten in der deutschen Politikwissenschaft. *Politische Vierteljahresschrift*, *48*(1), 97–117. https://doi.org/10.1007/s11615-007-0008-7

Podolny, J. (2005). *Status Signals*. Princeton University Press.

Posner, E. (2000). *Law and Social Norms*. Harvard University Press.

Rosenbaum, P. (2010). *Design of observational studies*. Springer.

Rossi, P., & Anderson, A. (1982). The factorial survey approach: An introduction. In P. Rossi & S. Nock (eds.), *Measuring Social Judgments: The Factorial Survey Approach* (pp. 1–25). Sage Publications.

Rubin, D. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, *2*(3), 808–840. https://doi.org/10.1214/08-AOAS187

Sanz-Menéndez, L., Cruz-Castro, L., & Alva, K. (2013). Time to Tenure in Spanish Universities: An Event History Analysis. *PLOS ONE*, *8*(10), e77028. https://doi.org/10.1371/journal.pone.0077028

Sauer, C., Auspurg, K., Hinz, T., & Liebig, S. (2011). The Application of Factorial Surveys in General Population Samples: The Effects of Respondent Age and Education on Response Times and Response Consistency. *Survey Research Methods*, *5*(3), 89–102. https://doi.org/10.18148/srm/2011.v5i3.4625

Searcy, A. & Nowicki, S. (2005). *The Evolution of Animal Communication. Reliability and Deception in Signalling Systems.* Princeton University Press.

Schröder, M., Lutter, M., & Habicht, I. (2021). Publishing, signaling, social capital, and gender: Determinants of becoming a tenured professor in German political science. *PLOS ONE*, *16*(1), e0243514. https://doi.org/10.1371/journal.pone.0243514

Schulze, G., Warning, S., & Wiermann, C. (2008). What and How Long Does It Take to Get Tenure? The Case of Economics and Business Administration in Austria, Germany and Switzerland. *German Economic Review*, *9*(4), 473–505. https://doi.org/10.1111/j.1468-0475.2008.00450.x

Simpson, A. (2017). The surprising persistence of Biglan's classification scheme. *Studies in Higher Education*, *42*(8), 1520–1531. https://doi.org/10.1080/03075079.2015.1111323

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (second edition). Sage Publications.

Solga, H., Rusconi, A., & Netz, N. (2023). Professors' gender biases in assessing applicants for professorships. *European Sociological Review*, jcad007. https://doi.org/10.1093/esr/jcad007

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Stiglitz, J. (1975). The Theory of "Screening", Education, and the Distribution of Income. *American Economic Review*, *65*(3), 283–300. https://www.jstor.org/stable/1804834

Verlegh, P., Schifferstein, H., & Wittink, D. (2002). Range and Number-of-Levels Effects in Derived and Stated Measures of Attribute Importance. *Marketing Letters, 13*(1): 41–52. https://doi.org/10.1023/A:1015063125062

Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, *46*(8), 1416–1436. https://doi.org/10.1016/j.respol.2017.06.006

Weisshaar, K. (2017). Publish and Perish? An Assessment of Gender Gaps in Promotion to Tenure in Academia. *Social Forces*, *96*(2), 529–560. https://doi.org/10.1093/sf/sox052

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817–838. https://doi.org/10.2307/1912934

Williams, W., & Ceci, S. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences, 112*(17), 5360–5365. https://doi.org/10.1073/pnas.1418878112

Appendix

Table A1 Correlations (r) of fictitious candidates' characteristics

Experimental Design	1	2	3	4	5	6	7	8	9	10
1 Qualification	1.000									
2 Publications (German)	-0.011	1.000								
3 Publications (international)	0.029	0.001	1.000							
4 Teaching experience	-0.000	0.008	-0.001	1.000						
5 Third-party funding	-0.005	-0.015	-0.017	-0.006	1.000					
6 International experience during the PhD	0.009	0.008	0.003	-0.010	0.032	1.000				
7 International experience during the postdoc	-0.016	-0.027	0.035	0.007	-0.004	-0.003	1.000			
8 International networks	0.007	0.012	0.003	0.011	-0.019	0.002	0.029	1.000		
9 Citations (German)	0.009	0.004	-0.014	-0.010	0.017	0.021	0.009	-0.025	1.000	
10 Citations (international)	0.003	-0.023	0.074	0.009	-0.021	0.005	0.012	0.028	0.005	1.000

Source: SciMo Survey of Professors (2018).

Table A2 Estimations of main effects of vignette dimensions and effects of interaction with professors' discipline

Dependent variable: Suitability for a full professorship	M1	M2
Qualification		
Junior professorship (evaluated) (ref. none)	0.999***	1.431***
	(14.826)	(11.148)
Habilitation (postdoctoral qualification) (ref. none)	1.417***	1.999***
	(19.818)	(14.213)
Non-tenured associate (W2) professorship (ref. none)	1.441***	2.012***
	(19.343)	(13.740)
High number of German publications (ref. low number)	0.343***	0.498***
	(8.474)	(6.388)
High number of international publications (ref. low number)	1.026***	0.647***
	(20.487)	(8.100)
Much teaching experience (ref. little)	1.155***	1.132***
	(23.720)	(13.087)
Much third-party funding (ref. little)	1.092***	0.833***
	(22.808)	(9.733)
PhD gained abroad (ref. in Germany)	0.0475	-0.0944
	(1.149)	(-1.221)
Postdoc gained abroad (ref. in Germany)	-0.109*	-0.269**
	(-2.347)	(-3.288)
Contact with many scientists abroad (ref. few)	0.449***	0.421***
	(10.145)	(5.279)
High number of citations in German publications (ref. low number)	0.419***	0.540***
	(10.446)	(7.114)
High number of citations in international publications (ref. low number)	0.982***	0.696***
	(19.986)	(8.682)
Discipline of responding professor		
Social sciences (ref. German studies)	0.293*	0.204
	(2.476)	(0.845)
Chemistry (ref. German studies)	-0.810***	-0.996***
	(-6.359)	(-4.442)
Interaction terms		
Qualification		
Social sciences * Junior professorship (evaluated)		-0.483**
		(-2.844)
Social sciences * Habilitation (postdoctoral qualification)		-0.683***
		(-3.677)
Social sciences * Non-tenured associate (W2) professorship		-0.737***
		(-3.776)
Chemistry * Junior professorship (evaluated)		-0.734***
		(-4.467)
Chemistry * Habilitation (postdoctoral qualification)		-0.958***

		(-5.469)
Chemistry * Non-tenured associate (W2) professorship		-0.887***
		(-4.838)
Social sciences * High number of German publications		-0.113
		(-1.114)
Chemistry * High number of German publications		-0.307**
		(-3.038)
Social sciences * High number of international publications		0.581***
		(4.770)
Chemistry * High number of international publications		0.520***
		(4.608)
Social sciences * Much teaching experience		-0.0274
		(-0.224)
Chemistry * Much teaching experience		0.0651
		(0.551)
Social sciences * Much third-party funding		0.181
		(1.497)
Chemistry * Much third-party funding		0.528***
		(4.666)
Social sciences * PhD gained abroad		0.174
		(1.645)
Chemistry * PhD gained abroad		0.222*
		(2.213)
Social sciences * Postdoc gained abroad		0.00362
		(0.032)
Chemistry * Postdoc gained abroad		0.420***
		(3.785)
Social sciences * Contact with many scientists abroad		0.0846
		(0.764)
Chemistry * Contact with many scientists abroad		-0.000951
		(-0.009)
Social sciences * High number of citations in German publications		-0.0833
		(-0.811)
Chemistry * High number of citations in German publications		-0.240*
		(-2.451)
Social sciences * High number of citations in international publications		0.346**
		(2.925)
Chemistry * High number of citations in international publications		0.461***
		(4.057)
Constant	0.900**	0.995**
	(2.672)	(2.809)
σ_u	1.127	1.121
σ_e	1.700	1.674
ρ	0.305	0.310
R^2	0.323	0.337

$R^2_{between}$	0.148	0.150
R^2_{within}	0.393	0.414
Wald χ^2	3181.6***	3550.6***
$N_{vignettes}$	6354	6354
$N_{respondents}$	874	874

Random effects regression with covariates at respondents' level (see section 3.2 for details)
b-coefficients, robust standard errors, *t* values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: SciMo Survey of Professors (2018).

Table A3 Robustness checks (estimations with covariates, fixed effects, and reduced sample)

Dependent variable: Suitability for a full professorship	M1	M2	M3	M4
Qualification				
Junior professorship (evaluated) (ref. none)	0.999***	0.999***	0.992***	1.022***
	(14.857)	(14.826)	(14.696)	(14.107)
Habilitation (postdoctoral qualification) (ref. none)	1.419***	1.417***	1.423***	1.436***
	(19.877)	(19.818)	(19.887)	(18.606)
Non-tenured associate (W2) professorship (ref. none)	1.441***	1.441***	1.439***	1.474***
	(19.377)	(19.343)	(19.251)	(18.261)
High number of German publications (ref. low number)	0.340***	0.343***	0.333***	0.351***
	(8.441)	(8.474)	(8.239)	(8.183)
High number of international publications (ref. low number)	1.024***	1.026***	1.027***	1.067***
	(20.502)	(20.487)	(20.450)	(19.836)
Much teaching experience (ref. little)	1.157***	1.155***	1.159***	1.192***
	(23.797)	(23.720)	(23.767)	(22.695)
Much third-party funding (ref. little)	1.091***	1.092***	1.090***	1.093***
	(22.872)	(22.808)	(22.735)	(21.086)
PhD gained abroad (ref. in Germany)	0.0473	0.0475	0.0505	0.0685
	(1.145)	(1.149)	(1.219)	(1.547)
Postdoc gained abroad (ref. in Germany)	-0.108*	-0.109*	-0.101*	-0.0750
	(-2.325)	(-2.347)	(-2.171)	(-1.489)
Contact with many scientists abroad (ref. few)	0.451***	0.449***	0.456***	0.453***
	(10.214)	(10.145)	(10.317)	(9.530)
High number of citations in German publications (ref. low number)	0.420***	0.419***	0.420***	0.418***
	(10.523)	(10.446)	(10.492)	(9.794)
High number of citations in international publications (ref. low number)	0.981***	0.982***	0.984***	1.047***
	(20.031)	(19.986)	(20.020)	(19.770)
Constant	1.052***	0.900**	1.053***	0.923***
	(11.197)	(2.672)	(13.219)	(9.233)
σ_u	1.222	1.127	1.430	1.181
σ_e	1.700	1.700	1.700	1.710
ρ	0.341	0.305	0.414	0.323
R^2	0.273	0.323	0.273	0.289
$R^2_{between}$	0.0285	0.148	0.0282	0.0100
R^2_{within}	0.393	0.393	0.393	0.404
Wald χ^2 / F	2828.6***	3181.6***	233.2***	2646.9***
$N_{vignettes}$	6354	6354	6354	5528
$N_{respondents}$	874	874	874	691

M1: Random effects regression without covariates
M2: Random effects regression with covariates at respondents' level (see section 3.2 for details)
M3: Fixed effects regression (Hausman test: $\chi^2 = 23.84$, p = 0.0214)
M4: Random effects regression without dropouts at respondents' level
b-coefficients, robust standard errors, *t* values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: SciMo Survey of Professors (2018).