

WORK AND LIFE IN ACADEMIA

Assessing educational expertise in academic faculty promotion

Sara Levander* and Ulla Riis

Department of Education, Uppsala University, Uppsala, Sweden

During 1999–2010, eligible Swedish university lecturers had an unconditional right to apply for promotion to the position of professor. Our aim was to discuss the motives of the reform and to problematise challenges in making qualitative assessments of educational expertise. We presented the results from an evaluation of the reform, and we focused on the weights that the peer reviewers in their assessment assign to the educational credentials of the applicants as opposed to those assigned to the research credentials. The empirical material consists of the dossiers from 294 cases of promotion. For research expertise and for educational expertise, we created one and three indices, respectively, where different types of credentials were given different weights. Changes over time were examined, as well as differences between disciplinary domains. In the assessment and decision process, educational expertise was outweighed by research expertise, and mainly quantitative aspects of the former were taken into account. There were signs that the peer review system underwent changes and that its intended quality-promoting function diminished over time.

Keywords: *academic work; educational expertise; evaluation; peer review; promotion*

*Correspondence to: Sara Levander, Department of Education, Uppsala University, Box 2136, S-750 02 Uppsala, Sweden. Email: sara.levander@edu.uu.se

During 1999–2010, eligible Swedish university lecturers had the unconditional right to apply to be promoted to the position of full professor. There were three official political motives: (1) increasing the proportion of professors in the faculty corps, (2) furthering teaching and educational work among university faculty and (3) furthering diversity, including gender, in university research and education (Proposition, 1996/97:141).

These motives can be assumed to impose threats to traditional academic (research) values.

The reform was considered a rather large one, and when it was launched several concerns arose, such as: Would promoted professors come to replace chaired professors appointed in competition after a public posting? Would the level of competence of promoted professors be on a par with that of chaired professors? Since no economic resources were added to the implementation of the reform, the academic labour unions and many lecturers aspiring to be promoted feared it would merely be a title reform.

Eligibility for promotion was defined thus: A person having ‘demonstrated *both* scientific qualifications and teaching qualifications shall be qualified for employment as a professor’ (Högskoleförordningen, 2009/2010, our emphasis). The word ‘both’ implied that both types of qualifications must be scrutinised, that they must

independently of each other reach the level demanded for a professor and that strength in one of them may compensate for weakness in the other only if this level is exceeded (Högskoleverket, 2007). Ahead of the reform, it was also emphasised that the importance of teaching qualifications, or rather educational expertise in a broader sense, should be considered to a greater extent than before. Previously, a professorship could be attained only in competition as the number of chairs was in principle fixed through decisions by the Swedish Parliament and Government. This emphasis on teaching qualifications – as compared with scientific qualifications – was preceded by recurrent expectations from society on Academia for some 40–50 years (Lindberg, 2005), expectations which had been largely neglected by the academic society, especially in peer review processes (SOU, 1996:166).

In 2011, an evaluation of this promotion reform was performed (Riis, Hartman, & Levander, 2011). We will present the results regarding the increase in the number of professors and the educational issues (motives 1 and 2), whereas we are dealing with diversity and gender (motive 3) in ongoing research.

In this article, we focus on questions pertaining to educational expertise and its corresponding credentials as well as on how these were judged and assessed in the promotion process. The information given by the applicants

for promotion along with the written statements of the peer reviewers¹ form our prime empirical source. Our aim is to account for a reform and its driving forces, as well as for some of the outcomes. Our aim is also to problematise the challenges of making qualitative assessments of educational expertise. These issues have heretofore received relatively little attention in research on the higher education system.

We argue that in this process of judging expertise, educational expertise is outweighed by research expertise and we also claim that mainly quantitative aspects are taken into account, especially regarding educational expertise. Furthermore, there are signs that the peer review system has undergone changes and that its quality-promoting function has diminished over time.

We will designate scientific qualifications *research expertise* and teaching qualifications *educational expertise*.² By 'expertise', we mean the value of a person's expert knowledge attributed to him or her by somebody else, in the case of academic promotion by the peer reviewers and, finally, by the decision-maker, in our case the Vice-Chancellor of the university. By 'credentials' – educational or research credentials – we mean the achievements regarding education and research that are allegedly objectively described by an applicant for promotion.

One originality of this article lies in the use of curricula vitae (CVs) and peer reviewers' written evaluations as a method for studying assessments for promotion, as well as the focus on the challenges of assessing educational expertise as part of an academic reward system.

What is known about academic reward systems?

O'Meara (2011) argued that academic reward systems form an ever-present system of participation, action and consequences that influence faculty priorities and careers. Rice, Sorcinelli, and Austin (2000) as well as Tierney and Bensimon (1996) pointed out that aspirations for tenure and for promotion, in particular, are highly influential in shaping the behaviours of early-career faculty. Any academic reward system, however organised, relies heavily upon peer review, that is, academics scrutinising and valuing the accomplishments of other academics.

Following Boyer's (1990) urge for a broader definition of *scholarship* and more flexible criteria for tenure, several universities in the United States have changed their systems to support multiple forms of scholarship (Leisyte & Dee, 2012). However, despite this, most universities continue to prioritise research in their faculty reward systems (Leisyte & Dee, 2012), and outcomes like promotion,

tenure, pay, etc. are primarily connected to faculties' performances in research (Braxton, Luckey, & Helland, 2002; Fairweather, 2005; Melguizo & Strober, 2007; O'Meara, 2011).

Assessment of research expertise is a complex phenomenon, often reduced to the use of standard procedures, such as quantities of number of citations or number of articles published in journals of certain reputations. Assessing, valuing and rewarding educational expertise can, however, also be considered a challenging task (Braxton et al., 2002; Burkill, 2002; Fairweather, 2005; Gibbs, 1995; Melguizo & Strober, 2007; Prosser, Benjamin, Martin, & Trigwell, 2000; Ramsden, Margetson, Martin, & Clarke, 1995; Ramsden & Martin, 1996; Trigwell, 2001). It would, therefore, not be surprising if educational expertise turns out to be an equally complex phenomenon and conjures up a similar reduction.

It has thus been shown that the academic reward system influences the behaviour of faculty, especially that of younger faculty, and that this influence involves applying a higher value to research activities than to educational ones. Given a choice, the typical academic will, therefore, prefer research over teaching and strive to amass research credentials rather than educational ones. This conceivably influences the quality of teaching and what and how students learn, and in the long term, the academic system itself.

What about the driving forces behind the preference of research over education? Some 80 years ago Ludwik Fleck, the Polish scientist and philosopher, launched the concepts 'thought collective' and 'thought style' (Fleck, 1935/1979), thereby anticipating later theories on paradigm (Kuhn, 1962). Peer reviewers, belonging to one and the same academic culture and 'thought collective', presumably share the same 'thought style'. In our case, the thought style harbours a tradition of seeing the largest academic value in research and of safeguarding academic freedom: the freedom of autonomously choosing one's research problem, the freedom of autonomously choosing one's research methods and the freedom of freely publishing one's findings. These 'freedoms' are laid down in the Swedish Law of Higher Education and discussed in general in the scientific literature by, for example, Aarrevaara (2010), Höhle and Teichler (2013) and Musselin (2013). Often the idea of 'academic freedom' also embraces the idea of freedom of choosing one's value criteria, pretensions that have been heavily criticised (Stockfelt, 1982). However, an open and meritocratic society is dependent upon peer review, and we hold this a premise.

Fairweather (2002,2005) pointed to the challenges in assessing academic work on the whole. He stipulates that assessing '*accountability* requires different measures and standards than do assessments of *quality* and *impact*' (Fairweather, 2002, p. 103, emphasis in original). Accountability refers to, for example, the number of courses taught

¹'Peer reviewers' are sometimes also referred to as 'external experts', 'subject specialists', 'expert reviewers', 'experts' or 'referees' in the literature as well as in the written peer review statements we have examined.

²Educational expertise is also often referred to as 'teaching excellence' or 'good teaching' or 'scholarship of teaching and learning' in the literature cited here.

while quality might deal with instructional approaches or student learning outcomes. These concepts may be useful in our analysing and understanding the outcome of our study.

A study of a distinct and perspicuous promotion reform in Swedish Academia may add to our understanding of a peer review system with origins in the late 19th century³ and still exerting an influence on today's individual academics as well as on Academia itself (Lindberg, 2005).

The Swedish academic career system

The Swedish university system is a unified one, including all public post-secondary education, which is traditional 'free' studies in science and humanities, as well as education of professionals, ranging from, for example, medicine, engineering and law to, for example, nursing and primary school teaching.

An academic career in Sweden today⁴ starts with the aspirant being accepted at a 4-year doctoral program with a dissertation as the goal. The PhD degree is necessary for the eligibility as a university lecturer. All professorships are tenured and with very few exceptions a lectureship also means tenure. In 2011, a so-called tenure track system was introduced to be used *ad libitum* by each university institution, but it has been used quite seldom.⁵ Up until 1998, the number of professorships was regulated on a national level. After this date, each university institution decides itself on the proportion of professorships among its academic staff (lecturers and professors). Professors and lecturers alike are expected to teach and do research. However, the resources allotted to professors' research is considerably larger than that allotted to lecturers' research. The teaching load for lecturers used to exceed that of professors, but after the 1999 reform this is no longer granted.

Research design

Context

In the evaluation study (Riis et al., 2011), we wanted first to systematise the knowledge of what is colloquially referred to as 'the bar': What qualifications were required for a senior lecturer to be promoted to a professor? Were the *à priori* concerns valid? Were there differences in how research expertise and educational expertise were described, judged and assessed over time and across disciplinary domains?

The context of our study was the Swedish university system and the case studied was that of Uppsala University. Uppsala University, founded in 1477, is one of the two largest universities in Sweden with a turnover of

10–12% of the country's expenditures on universities (university colleges not included) (Högskoleverket, 2011).

The promotion reform, mandatory for the university institutions, was in operation 1999–2010. This period had been preceded by a very large expansion of higher education throughout the 1990s; the number of university students increased by 85% from 1991 to 2000 (SOU, 2001:13).

The political motive to increase the proportion of professors in the faculty corps was reached indeed: at the national level, there were 2,310 chaired professors in 1999 and at the end of 2010 the total number of professors, chaired and promoted, was some 4,515, of whom some 2,500 were promoted (Högskoleverket, 2000, 2011). Of these, some 20% were employed at Uppsala University. During the 12 years, there was a general expansion of the university system and the number of university teachers of all categories rose by 12%. Thus, the 95% increase in the number of professors was far greater than the general expansion.

It should also be noted that the teaching load and other working conditions were not 'automatically changed' (RALS, 1998) for a lecturer promoted to a professor. In practice, such a person would normally be assigned the same amount of teaching duties as before and would not have her time for research increased.⁶ She would receive a modest raise in salary, however.

In the 12-year period at Uppsala University, close to 700 senior lecturers applied for promotion and just over 500 were promoted. At the outset of the reform, the number of chaired professors was 294; by 2010, their number was 245, to be compared with 300 promoted professors.⁷ Early in 1999, the proportion of lecturers among faculty at Uppsala University was 72% and at the end of 2010 it was 50% (Uppsala universitet, 1999, Uppsala universitet, 2011).

The process of promotion in Sweden takes place in four steps, consisting of (1) the application by the person aspiring to be promoted. The following steps represent successive action: (2) the peer reviewers' recommendation, (3) the faculty board's suggestion and (4) the Vice-Chancellor's decision to promote or not promote. The peer reviewers' written statements are intended to guide the drafting of the matter as well as the Vice-Chancellor's decision.

Material

The empirical material included the dossiers from the promotion process: (1) applications, including CVs, from 294 (43%) of the 678 promotion cases at Uppsala University,⁸

⁶In the following, we will use the preposition 'her' instead of 'her or his'.

⁷The difference between 500 and 300 is due to the fact that during the 12-year period a large number of promoted professors retired.

⁸Due to the time frame of the project, it was not possible to work with all 678 cases of promotion why we set out for a sample of circa 33%, evenly spread over time and scientific domains. However, for practical reasons it was yet possible to include some 43% of the cases in the sample and we chose to do so.

³Today's Swedish academic peer review system can be traced back to a state ordinance from 1876 (Svensk författningssamling, 1876:5).

⁴Since approximately 1970.

⁵A reason may be that according to general Swedish labour market legislation in operation since the 1970s practically, all faculty are permanently employed.

(2) the peer reviewers' written statements – normally two, sometimes three per case, (3) the faculty board protocols and (4) the Vice-Chancellor's protocols. The written application is expected to contain a CV including a list of publications. Normally, it also comprises the applicant's presentation of her various career steps and experience of research and teaching. Examples of student evaluations may be included. The applicant also oftentimes includes a qualitative account of her personal educational philosophy, often prompted by instructions and templates issued by the university, addressing applicants as well as peer reviewers.⁹ These instructions stress the importance of educational expertise. The experiences accounted for in this way constitute the applicant's credentials. The peer reviewers are expected to scrutinise these credentials, including the presentation of the applicant's educational philosophy, if included, and based on this, to formulate an assessment of the applicant's research expertise and educational expertise, and to recommend that the applicant be, or not be promoted.

Methodology

The creation of indices

A recurrent issue of debate in Academia is whether the assessment of academic credentials is comparable over disciplinary domains. Are there systematic differences in the valuations of credentials regarding, for example, medicine versus humanities? Questions like this were augmented at the onset of the reform. We considered them relevant for our evaluation, and we therefore needed a method for comparison. For the research expertise and the educational expertise, we created one and three indices, respectively. The index for research expertise was based on 14 variables, each corresponding to a certain type of publication. It built on the number of publications and a relative weight assigned to each one of them, with a monograph for instance being assigned the weight of 12, whereas an article in an international journal was assigned a weight between 3 and 1, depending on the prominence of the applicant among the co-author(s), if there were any. We created three indices for educational expertise: *experience of teaching and educational planning* (TEP); *supervision of doctoral candidates* (SR); and *reflection about teaching issues* (REFLEC). TEP and SR were based on the number of quantitative credentials such as the amount of teaching in front of a class and the number of doctoral students supervised. Here, too, different types of credentials were given different weights: experience as director of studies, for example, weighing twice that of experience as director of doctoral studies. The third index, REFLEC, was aimed at capturing qualitative aspects

of educational expertise by noting, for instance, training in tertiary-level teaching, educational development work, etc. Table 1 illustrates 9 (out of 17) variables and corresponding composition of indices regarding educational expertise. The detailed method of identifying variables and of weighting and justification of weights is explained in our two previously published reports (Riis, 2012; Riis et al., 2011).

Through the weighting, we managed to calibrate three comprehensive disciplinary domains against each other, allowing them to assume approximately the same mean values. The three domains, corresponding to the overall organisation of Uppsala University, were: Humanities and Social Sciences (HS), Medicine and Pharmacy (MP) and Science and Technology (ST). Our indices have allowed comparisons over time, by disciplinary domains and by gender. In this article, the aspect of gender has been left aside.

The value of CVs and peer evaluations for studying academic reward systems

When Dietz, Chompalov, Bozeman, O'Neil Lane, and Park (2000) used CVs in their study of Academia, this was a data source out of the ordinary, but over the years it has become more frequent (Cañibano, Otamendi, & Solís, 2011). Cañibano and Bozeman (2009) pointed to three research areas in which CVs are used with advantage, among which are studies of researchers' career trajectories (e.g. Dietz et al., 2000; Dietz & Bozeman, 2005; Fukuzawa, 2014; Gaughan & Bozeman, 2002; Sabatier, Carrere, & Mangematin, 2006) and studies of mapping of collective capacity (e.g. Jonkers & Tjissen, 2008; Lee & Bozeman, 2005; Lin & Bozeman, 2006). Our study relates to studies of career trajectories. It is similar to that of Sabatier et al. (2006) who analysed factors influencing the length of time to promotion. Sabatier et al. (2006) and Sandström, Wold, Jordansson, Ohlsson, and Sandberg (2010) included publications in their data collections, and so did we.

Cañibano and Bozeman (2009) stated that although CVs provide a useful source of data, several methodological problems are attached to their use, such as availability, truncation, missing information and heterogeneity. Availability and truncation were not obstacles in our study, as the principle of public access to official records¹⁰ is applied in Sweden and the dossiers in question are available either at the local archive of each university or at the National Archives. Missing data were a minor problem, which tended only to affect information regarding fund raising, mobility and a few other aspects which we do not deal with here. Heterogeneity was somewhat of a problem regarding the first couple of years in the

⁹Uppsala University issued two booklets, one addressing applicants on how to structure a CV and one addressing peer reviewers on how to formulate expert statements (Apelgren & Giertz, 2002; Giertz, 2004).

¹⁰The Swedish principle of public access to official records has been in operation since the year 1766 and is today part of the Constitution (Svensk författningssamling, 1949:105).

Table 1. Examples of indices regarding *educational expertise*. Values normalised so that: $0 < x \leq 100$.

Variable – designation	Variable – definition	Included in index
Main supervisor, completed Weight 8	Number of doctoral candidates ushered to degree completion as main supervisor	SR
Main supervisor, ongoing Weight 4	Number of doctoral candidates with work in progress as main supervisor	SR
Director of studies basic level Weight 10	The applicant has served as director of studies for first- and second-cycle (basic and advanced undergraduate) education	TEP
Director of studies research level Weight 5	The applicant has served as director of studies for third-cycle (research level) education	TEP
In-service training Weight 1	The applicant has completed one or more short courses (a few days to a week) in tertiary-level teacher training	REFLEC
Publications on teaching issues Weight 2	Number of times the applicant has published on teaching issues	REFLEC
Textbooks Weight 6	Number of times the applicant has published a textbook	REFLEC
Tertiary-level teacher training Varying Weight	Own tertiary-level teacher training	REFLEC
Prizes and awards Weight 2	Number of distinguished-teaching prizes or awards won	REFLEC

period studied, after which the CVs tended to take on a more or less ‘standardised’ format, owing to recommendations from university agencies. Part of the problem of heterogeneity could also be approached in the coding process, which leads us to the problems of coding.

Cañibano and Bozeman (2009) and Dietz et al. (2000) have pointed to the large amount of work necessary to the coding of CVs. It is tedious and risks error due to coder fatigue and imperfect intercoder reliability (Corley, Bozeman, & Gaughan, 2003; Dietz et al., 2000). To avoid fatigue, we spread the coding over time and distributed it among the three of us. To counteract differences in interpretations, we continuously discussed the coding out from examples. Intercoder reliability was checked by cross-coding, especially in the early stages of coding, in order to achieve good alignment.

Peer reviewers’ written statements constituted an important component in the execution of the promotion reform (as it does in the recruitment system). The function of these evaluation statements is to ensure unbiased recruitments and high-quality appointments. In the Swedish literature, several studies have made use of such statements, probably thanks to the open access to public records (Gunvik-Grönbladh, 2014; Gunvik-Grönbladh & Giertz, 1998; Lindberg, 1988, 2005; Nilsson, 2009; Riis & Lindberg, 1996; Sandström et al., 2010). These authors constructed systematic methods for analysis of texts. In this article, we account for parts of the quantitative analysis in the evaluation by Riis et al. (2011). We account for results from a qualitative analysis on similar empirical data in ongoing research.

Results

Attention to research credentials prevails

As expected, research expertise measured via the index had a clear impact on the decision of whether or not to promote (cf. Braxton et al., 2002; Fairweather, 2005; Leisyte & Dee, 2012). Also, the applicants’ accounts of their credentials improved in structure and content over time. One prominent result was that the bar for research expertise was lowered across the 12 years and the bar for educational expertise was raised, in both cases though moderately. Table 2 shows the SR index.

When instead we regarded indices TEP and REFLEC, the results were not consistent with expectations, especially not for index REFLEC. Out of six cases (3 domains \times 2 quantitative indices) three went against expectations in which the mean score for promotion fell below that for denials (not shown in table). These partly unexpected findings lead us to the results regarding educational expertise, related to the final decisions by the Vice-Chancellor.

First, the overall importance of educational expertise for promotion rose over time – though not ‘at the expense’ of research expertise (see below). At this point, we might actually ask whether the Vice-Chancellor always made the ‘correct’ decision. Our indices could also be used to study the Vice-Chancellor’s decisions and Figure 1 illustrates this.

According to our indices, the Vice-Chancellor seems to have gone astray in deciding to deny promotion; these applicants score high in both research expertise and educational expertise. One of the two cases belongs to the MP domain and can be explained by the faculty board withholding information from the Vice-Chancellor

Table 2. Educational expertise, index SR and experience of PhD supervision by disciplinary domain and decision. Mean, N and SD.

Disciplinary domain	Decision	Mean	N	SD
Humanities and Social Sciences	Promoted	15.2	109	13.4
	Denial	5.4	22	7.8
	Total	13.6	131	13.1
Medicine and Pharmacy	Promoted	27.9	69	16.4
	Denial	19.1	7	6.4
	Total	27.1	76	15.9
Science and Technology	Promoted	23.8	74	15.4
	Denial	15.4	13	12.6
	Total	22.6	87	15.2

(see below). In three cases (2.5%), the Vice-Chancellor decided not to promote applicants with good educational credentials but with poor research credentials, but in 11 cases (9%), he decided against promotion for applicants having shown good research expertise but poor educational expertise. Regarding the Vice-Chancellor’s decision to promote, the pattern is not as clear; around 9% of those promoted had shown good educational

credentials but poor research credentials; for close to 8% the balance of credentials was the opposite.

Second, we saw that educational expertise, measured by the two quantitative indices, progressed in the expected direction. The more quantitative experience an applicant had as a university teacher, the greater her chance for promotion. Parts of the MP domain and the ST domain, however, diverged from the overall pattern. A certain number of doctoral students were regularly used as a minimum criterion of educational expertise, even though the Swedish Higher Education Appeals Board, in a precedential case, had rejected such a procedure (Högskoleverket, 2007). This minimum – and quantitative – criterion even evolved as a national practice within the domains of MP and ST.

Furthermore, the MP domain also deviated in their recommendations to the Vice-Chancellor, by apparently omitting information about strong educational credentials when these credentials seemed to further the chances for an applicant with poor research expertise.

Third, in contrast with the two quantitative indices, our qualitative index for educational expertise, REFLEC, neither furthered nor obstructed promotion.

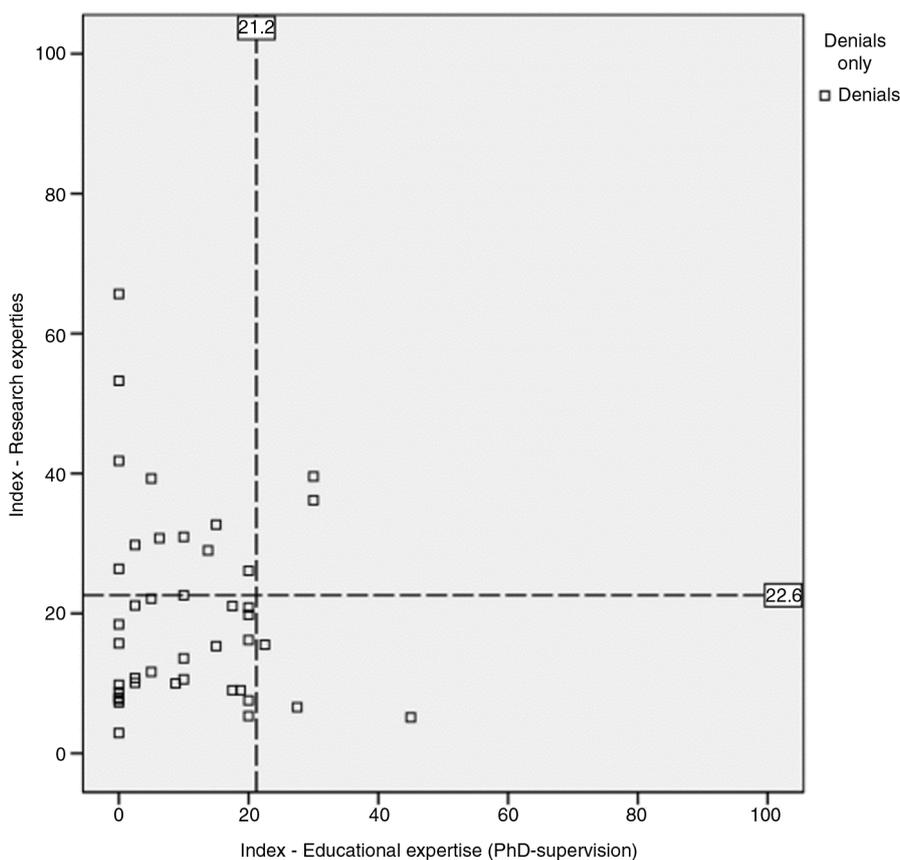


Fig. 1. Correlation between index for scientific credentials (research expertise) and index for teaching credentials (educational expertise regarding supervising) for those who were not promoted. Lines of demarcation indicate the means for the group promoted.

Thus, an applicant who tried to improve her skills as a teacher by participating in in-service training, initiating pedagogical projects or publishing about educational issues, could not count on promotion on that specific ground.

Notwithstanding, the overall result showed that research expertise was decisive in a clear majority of the cases. This could appear to be a contradiction, but even though educational expertise has increased in importance over time, research expertise was nevertheless the most influential criterion for the peer reviewers' recommendation as well as for the final decision.

What if we compare the Vice-Chancellor's decision with the peer reviewers' statements? Do the latter only serve as a 'soft' recommendation to the faculty board and the Vice-Chancellor? No, on the contrary, the Vice-Chancellor departs from the recommendations in these statements in only 3% of the cases, which includes the cases where the MP faculty board withheld information from him. The statements of the peer reviewers is obviously the prime source of information for the Vice-Chancellor in making his decision, as is the case, although to a lesser degree, for the faculty board. We conclude that peer reviewers' statements have a decisive role for the outcome of the promotion cases.

Changes in the work of the peer reviewers?

The peer reviewers are to be recruited from a higher education institution other than the applicant's. This is not only to avoid bias but also to maintain a national standard – 'the bar' in everyday language. The formalisation of the idea of a national standard goes some 100 years back, when 'the principle of inter-academic co-operation' was formally laid down in the national statutes for universities (Svensk författningssamling, 1916:66).

As mentioned, the applicants' descriptions of their own educational credentials improved over time in terms of quantity of information given, structure and so on. This improvement, however, did not prompt a corresponding effect in the peer reviewers' evaluations. Quite contrary, these statements were short, they became shorter over time, and they were most often void of argumentative information. Figure 2 illustrates the development of the length of the paragraphs regarding research expertise and educational expertise over the 12-year period.

Within the HS domain, the statements regarding research expertise clearly grew shorter with time. The statements regarding educational expertise remained unchanged within all three domains, probably, since the statements could hardly be further reduced – 5–10 cm corresponds to much less than half a page. The peer reviewers' statements were not

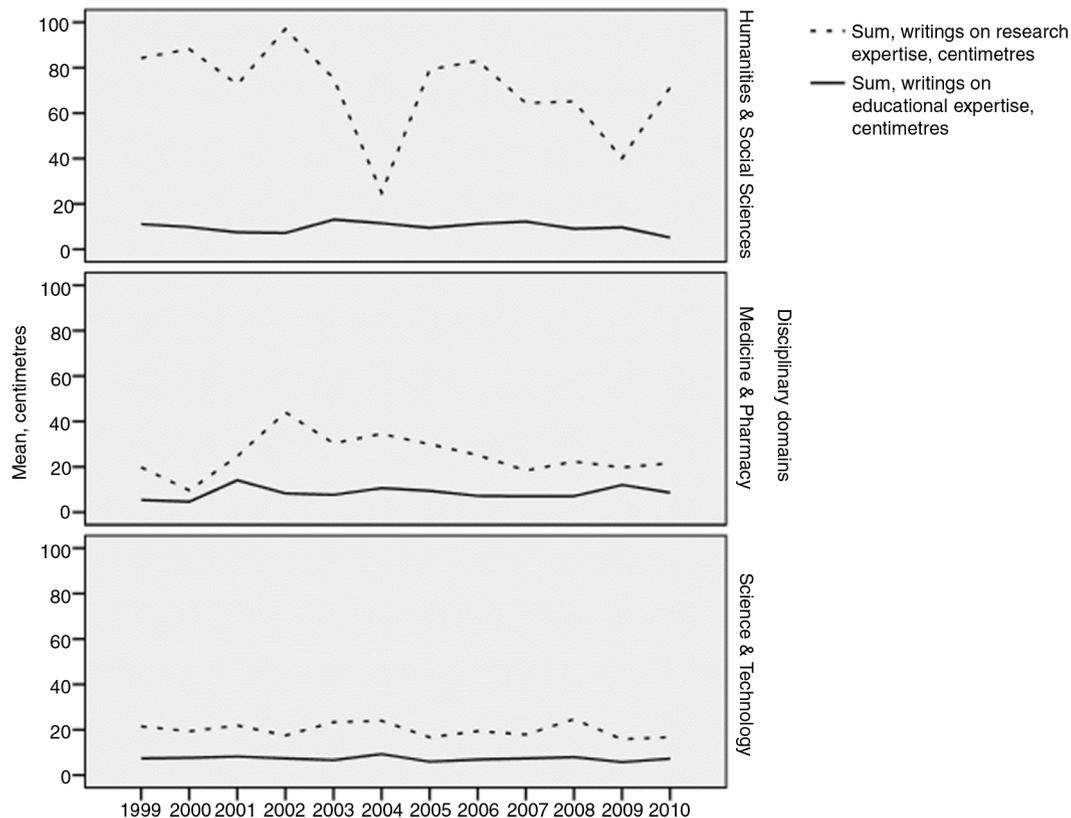


Fig. 2. The length in centimetres of the paragraphs regarding research expertise and educational expertise over the 12-year period; HS domain, MP domain, and ST domain. Mean value in centimetres.

only very short but also more often than not they also lacked feedback information. This occurred concurrently with the creation of instructions and support addressed not only to the applicants but also to the peer reviewers as well (see Note 9).

In her investigation, Nilsson (2009) also found very brief evaluation texts delivered by the peer reviewers, prompting her to call attention to threats to the quality of decision-making and to the legal security of the process. She hypothesised that impairments on the part of the task of the peer reviewers might be due to efforts to save time and money.

We conclude that in the execution of the promotion reform, the peer review system deteriorated, probably as a process and certainly regarding the deliverances.

Discussion

Our evaluation study was to follow up on two political motives for the Swedish promotion reform: (1) whether the proportion of professors in the corps of university faculty would increase and (2) whether educational achievements among faculty would be furthered by and rewarded in the processes of promotion. The first question has received an unequivocally affirmative answer.

The second question is more complicated and the overall answer is 'yes, but only partially and with some ignorance and reluctance'. The value attached to educational credentials in the Vice-Chancellor's decisions increased over the 12-year period, and more applicants with relatively good research expertise and relatively weak educational expertise were denied promotion than the reverse. Even so, the research credentials continued to be given priority, in the peer reviewers' statements as well as in the decisions by the faculty boards and by the Vice-Chancellor (cf. Fairweather, 2005; O'Meara, 2011). What can explain this finding? We see at least five circumstances in operation, and probably also interacting: 'rationalisation' of a demanding and partially tedious review work; imperfect documentation; tradition; perceived threats to academic freedom; lack of knowledge in combination with self-sufficiency.

Methodological questions

Before discussing the results, the question of representativeness should be addressed. Can results emanating from the case of Uppsala University be generalised? We think they can. Uppsala University is not only an old and large Swedish university with all traditional disciplines represented (except odontology), thus accommodating traditional and relatively conservative academic values.

However, and more important, in recruitment matters all higher education institutions in Sweden must adhere to the principle of interacademic cooperation laid down in 1876 (Universitetsstatuter, 1876). At the time, it was obligatory that the peers be recruited from another

university than the one responsible for the promotion (or the recruitment of a teacher after a public posting). This arrangement secured that national standards were met and that local interests were blocked.¹¹ In our view, the standards were upheld in the promotion processes studied, but as we have shown the standard as such varies with scientific domain and discipline. This finding is in accordance with the well-known findings, for example, Becher and Trowler (2001) and Lamont (2009) demonstrating quite large differences between disciplines regarding traditions and values but cohesiveness within each discipline. We believe that the differences present in our material are conditioned by disciplines and not by institutions, and that results regarding Uppsala University can be generalised to other (large) higher education institutions.

Rationalisation and the impact of supervising

The peer reviewers made adaptations to their way of evaluating various criteria of eligibility in order, we believe, to manage the task of assessing a large and increasing number of applications. This is especially conspicuous regarding evaluation of educational expertise, where supervising of doctoral students clearly evolved as a criterion and a manageable way to assess educational expertise.

Supervising doctoral students is a demanding task and requires both skills and engagement on the part of the supervisor. Today, thanks to regulations regarding the doctoral education, most of its students graduate irrespective of such skills or engagement on part of the supervisors. The criterion can thus be considered as nothing but a quantitative measure and we conclude that this criterion has been used in a routine manner. This criticism applies to the Vice-Chancellor, faculty boards and peer reviewers alike, even though the peer reviewers hold the crucial role in supplying the subsequent links of the chain of assessment by means of their written statements. Moreover, the criticism applies to a greater extent to the peer reviewers and the faculty boards of MP and of ST than to those of HS. We conclude that the Vice-Chancellor, on his part, clearly used the promotion reform to reward senior lecturers for good performances, especially within education.

Should supervising doctoral students be viewed as part of research expertise or educational expertise? In the natural sciences, the supervisor is often acknowledged as a co-author of the doctoral student's articles, and thus an applicant who has supervised such a student might choose whether to consider her contribution as a research credential or an educational credential. If she has strong research credentials and weak educational credentials, she could choose to enter supervising as an educational

¹¹Beginning in 2011, this national regulation was cancelled and it is now a matter for each university to include it, or not, in its local regulations. Practically, all universities have chosen to do so and thus to adhere to a long tradition, cf. Note 3.

credential. Such a choice stands out in many cases within the domains of MP and ST, but not of HS, and this exemplifies another conclusion, namely the fact that the variation between the different domains and different groups of disciplines is substantial, which may point to a need for further research on less aggregated data.

Our TEP index proved to be advantageous in the final decision in a similar way as that of supervising, but to a much lower degree and only in the HS and the ST domains. Thus, only supervising of doctoral students held a pervasive force in the final decision process. In Fairweather's (2002) terms, it might be easier to develop measures of accountability than measures of quality and impact.

Imperfect documentation: educational expertise – 'hard to find'?

One of our conclusions is that qualitative aspects of educational expertise are difficult to describe and assess. As presented above, our REFLEC index – intended to catch the core of professionalism of university teachers – did neither further nor obstruct promotion. Possible reasons for this might be imperfect documentation and/or peer reviewers not really knowing what or how to assess.

Is it at all possible to produce a fully comprehensible base for assessing educational expertise? The core of educational expertise is not directly observable in the manner that its counterpart is recognisable in research publications. This becomes even more obvious when taking into account the fact that the peer reviewers are supposed to be recruited from another university than the applicant's and generally lack direct knowledge of the applicant in question. Moreover, educational expertise is not normally reviewed by an evaluation prior to the peer reviewers' assessments, as research expertise has been, by being scrutinised in seminars or by journal reviewers. We have seen indications, however, that in cases where the applicant included a personal, written account of her educational philosophy, this is on occasion – but only on occasion – used by the reviewers as an approximation of the quality of the applicant's educational repertoire. This result reinforces the idea of accountability (Fairweather, 2002) being at stake rather than quality and impact. There is room for development here, both in practice and in scientific studies on the issue.

Tradition, academic freedom and self-sufficiency

Another explanation for the expert review system not having developed skills in assessing educational credentials lies with tradition (Nilsson, 2009). Expert reviewers belong to an academic culture and a 'thought collective' (Fleck, 1935/1979), and they presumably share the same 'thought style'. Inherent in such a thought style lies the professionally motivated desire to protect academic freedom. Societal pressures and official decrees, like the obligation to assess educational credentials as a prerequisite

for promotion, may be experienced as threats to both research and academic freedom.

Finally, as a professor one might intuitively think one 'knows' the level of educational expertise needed in one's own discipline (cf. Becher & Trowler, 2001). In combination with lack of knowledge of what to assess regarding educational expertise, this prompts ignorance to educational credentials. Thus, submission to policy documents and local guidelines may be regarded as a deviation from the common 'thought style' of the discipline.

Taken together, these circumstances may counteract external pressures. To the extent that the peer reviewers hold on to circumstances such as those discussed here, the high external expectations on them will lack legitimacy in their eyes. Herein lies a source of conflict between academic traditions and values on the one hand and the regulation, management and leadership of an expanding university system on the other.

A quality-driving system at risk?

One purpose of peer review is to maintain a national standard and to guarantee high quality in the assessments. Rather than conflicting interests between external agents and 'peers' this is at stake here. Peer reviewing, as part of a system, should contribute to the quality of research and teaching. If neither research credentials nor educational credentials are paid much – or any – particular attention by peer reviewers, one could call the function of the system into question. We consider the quite short statements to be deficient in quantity, structure and quality, and neither adequate nor practicable for quality-promoting work on the part of the individual applicant or the higher education system. As demonstrated, this was particularly true regarding assessment of educational expertise. This is despite the creation of support for the peer reviewers (see Note 9), and despite the fact that applicants have improved their way of describing their credentials, research credentials as well as the educational ones. We conclude that there is a resistance imbedded in the peer review system against external forces of change. Such forces include expectations from university leadership and university administration in times of expansion of higher education.

One can expect a promotion reform like this one to have an effect on the system level as well as on the individual level. If we regard it as part of a reward system, we expect the reform to influence academics' behaviour (Rice et al., 2000). This could be observed in the emerging ways research credentials and educational credentials were gathered, structured and presented by the applicants. In the longer perspective, the quality promotion function of the peer review system may be at risk. We find Fairweather's (2002) request for policy discussions concerning different kinds of faculty evaluations interesting

and relevant to prevent such a development in the Swedish academic system.

An evaluation study – and then?

The two political aims for the promotion reform which we addressed were reached: the number of professors increased, and educational credentials were assigned (a little) more weight over the 12 years studied. To lecturers and professors, the latter has probably signalled an increased importance in educational activities, though not of any great magnitude. The reform may be seen in an incremental light. A title reform was added to the leadership's arsenal of rewards.

We have also shown that peer reviewers hold a key position in the process of promotion. In their written evaluations, the peer reviewers clearly attached more weight to research expertise than to educational expertise. We have also seen that assessing educational expertise holds challenges and that several circumstances make it a difficult task. However, the study does not say much about what peer reviewers consider as constituting educational expertise, if they consider this at all.

In this article, we have primarily applied a quantitative approach, which has allowed analyses of a large-scale corpus. This approach does not, however, allow in-depth analyses of a more qualitative nature. Thus, the next research step will be to conduct textual analysis of written evaluations dealing with educational expertise in recruitments to professor chairs after a public posting. One might hypothesise that the assessments made in relation to advertised chairs hold a different content and formula – and hopefully quality – compared with those created in the process of promoting a local university lecturer to a professor.

Universities worldwide originate from different traditions and serve different societies. Despite contextual diversity of missions and practice, our results probably hold relevant for academic peer review activities, and the results from this Swedish case may contribute to an increased understanding of peer review in general and assessment of educational expertise in particular. In addition, our findings may serve as a base for comparative studies.

References

- Aarvevaara, T. (2010). Academic freedom in a changing academic world. *European Review*, 18, S55–S69. doi: <http://dx.doi.org/10.1017/S1062798709990317>
- Apelgren, K., & Giertz, B. (2002). *Skaffa dig en pedagogisk meritportfölj!* [Get yourself a teaching portfolio!]. Uppsala: Office for Development of Teaching and Interactive Learning (UPI), Uppsala University.
- Becher, T., & Trowler, P.R. (2001). *Academic tribes and territories. Intellectual inquiry and the culture of disciplines*. Buckingham: Open University Press.
- Boyer, E.L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.
- Braxton, J.M., Luckey, W., & Helland, P. (2002). The four domains of scholarship: Toward a rethinking of scholarly role performance. *ASHE-ERIC Higher Education Report*, 29(2), 11–26.
- Burkill, S. (2002). Recognising and rewarding excellent teachers: Towards a strategy for geography departments. *Journal of Geography in Higher Education*, 26(3), 253–262. doi: <http://dx.doi.org/10.1080/0309826022000019846>
- Cañibano, C., & Bozeman, B. (2009). Curriculum vitae method in science policy and research evaluation: The state-of-the-art. *Research Evaluation*, 18(2), 86–94. doi: <http://dx.doi.org/10.3152/095820209X441754>
- Cañibano, C., Otamendi, F.J., & Solís F. (2011). International temporary mobility of researchers: A cross-discipline study. *Scientometrics*, 89, 653–675. doi: <http://dx.doi.org/10.1007/s11192-011-0462-2>
- Corley, E., Bozeman, B., & Gaughan, M. (2003). Evaluating the impacts of grants on women scientists' careers: The curriculum vitae as a tool for research assessment. In P. Shapira & S. Kuhlmann (Eds.), *Learning from science and technology policy evaluation: Experiences from the U.S. and Europe* (pp. 293–315). Cheltenham, UK: Edward Elgar.
- Dietz, J.S., & Bozeman, B. (2005). Academic careers, patents and productivity: Industry experience as scientific and technical human capital. *Research Policy*, 34(3), 349–367. doi: <http://dx.doi.org/10.1016/j.respol.2005.01.008>
- Dietz, J.S., Chompalov, I., Bozeman, B., O'Neil Lane, E., & Park, J. (2000). Using the curriculum vitae to study the career paths of scientists and engineers: An exploratory assessment. *Scientometrics*, 49(3), 419–442. doi: <http://dx.doi.org/10.1023/A:1010537606969>
- Fairweather, J.S. (2002). The ultimate faculty evaluation: Promotion and tenure decisions. *New Directions for Institutional Research*, 2002, 97–108. doi: <http://dx.doi.org/10.1002/ir.50>
- Fairweather, J.S. (2005). Beyond the rhetoric: Trends in the relative value of teaching and research in faculty salaries. *Journal of Higher Education*, 76(4), 401–422.
- Fleck, L. (1935/1979). *Genesis and development of a scientific fact*. Chicago, IL: University of Chicago Press.
- Fukuzawa, N. (2014). An empirical analysis of the relationship between individual characteristics and research productivity. *Scientometrics*, 99(3), 785–809. doi: <http://dx.doi.org/10.1007/s11192-013-1213-3>
- Gaughan, M., & Bozeman, B. (2002). Using curriculum vitae to compare some impacts of NSF research grants with research center funding. *Research Evaluation*, 11(1), 17–26. doi: <http://dx.doi.org/10.3152/147154402781776952>
- Gibbs, G. (1995). Promoting excellent teaching is harder than you'd think: A note from an outside observer of the roles and rewards initiative. *Change: The Magazine of Higher Learning*, 27(3), 16–20.
- Giertz, B. (2004). *Assessing teaching skills in higher education*. Uppsala: Office for Development of Teaching and Interactive Learning (UPI), Uppsala University.
- Gunvik-Grönbladh, I. (2014). *Att bli bemött och att bemöta. En studie om meritering i tillsättning av lektorat vid Uppsala universitet* [Considering and being considered – A study of qualifications used in an appointment process of assistant professors at Uppsala University] (PhD dissertation, University of Uppsala, Sweden).
- Gunvik-Grönbladh, I., & Giertz, B. (1998). *Pedagogisk och Vetenskaplig Skicklighet i lika män? En Kartläggning av Pedagogiska*

- Meriter vid Tillsättning av Lektorat* [Educational expertise and research expertise – Do they count equal?]. Uppsala: Office for Development and Evaluation, Uppsala University.
- Högskoleförordningen. (2009/10). [Higher education ordinance]. Stockholm: Ministry of Education and Research.
- Högskoleverket. (2000). *Årsrapport för universitet och högskolor 1999* [National agency for higher education: Annual report regarding universities and university colleges 1999]. Stockholm: Swedish National Agency for Higher Education.
- Högskoleverket. (2007). *Befordran till professor och lektor – en rättslig översikt* [Promotion to professor and to senior lecturer – A legal overview]. Högskoleverkets Rapportserie, 55 R. Stockholm: Swedish National Agency for Higher Education.
- Högskoleverket. (2011). *Universitet och högskolor. Högskoleverkets årsrapport 2011* [Swedish National Agency for Higher Education: Annual report regarding universities and university colleges 2011]. Stockholm: Swedish National Agency for Higher Education.
- Höhle, E.A., & Teichler, U. (2013). *The work situation of the academic profession in Europe: Findings of a survey in twelve countries*. New York: Springer. doi: <http://dx.doi.org/10.1007/978-94-007-5977-0>
- Jonkers, K., & Tjissen, R. (2008). Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. *Scientometrics*, 77(2), 309–333. doi: <http://dx.doi.org/10.1007/s11192-007-1971-x>
- Kuhn, T.S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment*. Cambridge, MA: Harvard University Press.
- Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673–702. doi: <http://dx.doi.org/10.1177/0306312705052359>
- Leisyte, L., & Dee, J.R. (2012). Understanding academic work in a changing institutional environment. Faculty autonomy, productivity and identity in Europe and the United States. In J. Smart & M. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 123–206). Dordrecht: Springer. doi: http://dx.doi.org/10.1007/978-94-007-2950-6_3
- Lin, M.-W., & Bozeman, B. (2006). Researchers' industry experience and productivity in university–industry research centers: A 'scientific and technical human capital' explanation. *The Journal of Technology Transfer*, 31(2), 269–290. doi: <http://dx.doi.org/10.1007/s10961-005-6111-2>
- Lindberg, L. (1988). *Sakkunnigutlåtanden vid tillsättningar av professorer i pedagogik 1910–1982* [Experts' evaluations in the recruitment of professors in education 1910–1982]. Arbetsrapporter från Pedagogiska institutionen nr 63. Umeå: Umeå University.
- Lindberg, L. (2005). *Professorstillsättningar. Sakkunniga, professorstillsättningar och professorer i pedagogik 1910–1998* [Professor appointments. Peer reviewers, professor appointments and Swedish professoriate in education 1910–1998]. Arbetsrapporter från Institutionen för pedagogik nr 6. Växjö: Växjö University.
- Melguizo, T., & Strober, M.H. (2007). Faculty salaries and the maximization of prestige. *Research in Higher Education*, 48(6), 633–668. doi: <http://dx.doi.org/10.1007/s11162-006-9045-0>
- Musselin, C. (2013). How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate. *Research Policy*, 42(5), 1165–1173. doi: <http://dx.doi.org/10.1016/j.respol.2013.02.002>
- Nilsson, R. (2009). God vetenskap: hur forskares vetenskapsuppfattningar uttryckta i sakkunnigutlåtanden förändras i tre skilda discipliner [Good science: How researchers' conceptions of science expressed in peer review documents change in three different disciplines] (PhD dissertation, University of Gothenburg). Göteborg: University of Gothenburg.
- O'Meara, K.A. (2011). Inside the panopticon: Studying academic reward systems. In J. Smart & M. Paulsen (Eds.), *Higher education: Handbook of theory and research* (pp. 161–220). Dordrecht: Springer. doi: http://dx.doi.org/10.1007/978-94-007-0702-3_5
- Proposition. (1996/97:141). *Högskolans ledning, lärare och organisation*. [Government proposition 1996/97:141. The management, the teachers and the organisation of Swedish Higher Education]. Stockholm: Ministry of Education and Research.
- Prosser, M., Benjamin, J., Martin, E., & Trigwell, K. (2000). Scholarship of teaching: A model. *Higher Education Research & Development*, 19, 155–168. doi: <http://dx.doi.org/10.1080/072943600445628>
- RALS. (1998). *Ramavtal om löner m.m. 1998-09-24* [Centrally reached agreement between the employer organisation and the employee organisations regarding salaries and other working conditions regarding the state universities]. Stockholm: The Swedish Agency for Government Employers.
- Ramsden, P., Margetson, D., Martin, E., & Clarke, S. (1995). *Recognising and rewarding good teaching in Australian universities*. Canberra, ACT: Australian Government Publishing Service.
- Ramsden, P., & Martin, E. (1996). Recognition of good university teaching: Policies from an Australian study. *Studies in Higher Education*, 21, 299–315. doi: <http://dx.doi.org/10.1080/03075079612331381231>
- Rice, R.E., Sorcinelli, M., & Austin, A. (2000). *Heeding new voices: Academic careers for a new generation*. New pathways working paper series, Inquiry #7. Washington, DC: American Association for Higher Education.
- Riis, U. (2012). *Is the bar quivering? What can we learn about academic career requirement from the 1999 promotion reform?* Pedagogisk Forskning i Uppsala 161. Uppsala: Uppsala University.
- Riis, U., Hartman, T., & Levander, S. (2011). *Darr på ribban? En uppföljning av 1999 års befordringsreform vid Uppsala universitet* [Is the bar quivering? A follow-up study at Uppsala University of the promotion reform of 1999]. Acta Universitatis Upsaliensis, Uppsala Studies in Education 127. Uppsala: Uppsala University.
- Riis, U., & Lindberg, L. (1996). *Värdering av Kvinnors respektive Mäns Meriter vid Tjänstetillsättning inom Universitet och Högskolor* [Assessment of female and male qualifications in appointments within higher education]. Ministry of Education and Research. Stockholm: Ministry of Education and Research.
- Sabatier, M., Carrere, M., & Mangematin, V. (2006). Profiles of academic activities and careers: Does gender matter? An analysis based on French life scientists' CVs. *Journal of Technology Transfer*, 31(3), 311–324. doi: <http://dx.doi.org/10.1007/s10961-006-7203-3>
- Sandström, U., Wold, A., Jordansson, B., Ohlsson, B., & Sandberg, Å. (2010). *Hans excellens. Om miljardsatsningarna på starka forskningsmiljöer* [His excellency: On the investments of billions in centers of excellence]. Stockholm: Swedish National Agency for Higher Education.
- SOU. (1996:166). *Lärare för högskola i utveckling* [Teachers for higher education in transition]. Stockholm: Fritzes.
- SOU. (2001:13). *Nya villkor för lärandet i den högre utbildningen* [New conditions for student learning in higher education]. Stockholm: Fritzes.
- Stockfelt, T. (1982). *Professorn* [The professor]. Stockholm: MaxiMedia.

- Svensk författningssamling. (1876:5). Kongl. Maj:ts förnyade och nådiga statuter för universiteten i Upsala och Lund. [The Royal Majesty's renewed and gracious statutes of the universities of Upsala and Lund] Stockholm: The King-in-Council.
- Svensk författningssamling. (1916:66). Kungl. Maj:ts nådiga statuter för universiteten i Uppsala och Lund. [The Royal Majesty's gracious statutes of the universities of Uppsala and Lund] Stockholm: The King-in-Council.
- Svensk författningssamling. (1949:105). Tryckfrihetsförordning. [Freedom of the press act] Stockholm: Ministry of Justice.
- Tierney, W., & Bensimon, E. (1996). *Promotion and tenure: Community and socialization in academe*. Albany, NY: State University of New York Press.
- Trigwell, K. (2001). Judging university teaching. *International Journal for Academic Development*, 6(1), 65–73. doi: <http://dx.doi.org/10.1080/13601440110033698>
- Uppsala universitet. (1999). *Årsredovisning 1998* [Annual report regarding 1998]. Uppsala: Uppsala University.
- Uppsala universitet. (2011). *Årsredovisning 2010* [Annual report regarding 2010]. Uppsala: Uppsala University.