DEPARTMENT OF EDUCATION AND SPECIAL EDUCATION

ELUSIVE QUALITY OF EDUCATIONAL RESEARCH IN A CONTEXT OF BIBLIOMETRICS-BASED RESEARCH FUNDING SYSTEMS

The Case of the University of Gothenburg 2005-2014

Linda Sīle

Master Thesis:	30 credits
Program:	International Master in Educational Research
Level:	Second cycle (Advanced)
Term/Year:	Spring 2016
Supervisors:	Sverker Lindblad ad Gustaf Nelhans
Examiner:	Elisabet Öhrn
Rapport nr:	VT16 IPS PDA184:23

Abstract

Master Thesis:	30 credits
Program:	International Master in Educational Research
Level:	Second cycle (Advanced)
Term/Year:	Spring 2016
Supervisors:	Sverker Lindblad ad Gustaf Nelhans
Examiner:	Elisabet Öhrn
Rapport nr:	VT16 IPS PDA184:23
Keywords:	bibliometrics; quality; performance-based research funding systems; educational research; social epistemology; Sweden

---

Aim: This study aims to explore and theorise changes in publishing patterns within educational research in a context where bibliometric indicators are used in calculations for research funding distribution purposes. In theorising these publishing patterns, I discuss how such changes might and might not relate to the quality of educational research.

Theory: To do so, I employ a theoretical lens that is derived from Steve Fuller's social epistemology and adapted to the inquired case – publishing patterns within educational research at the Faculty of Education of the University of Gothenburg (Sweden) in a context of institutional bibliometrics-based research funding system.

Method: The empirical part of the study employs bibliometrics to explore change in publishing patterns. The time frame of the study is a period of ten years (2005-2014). In 2009, a performance-based research funding system (PRFS) was introduced at the University of Gothenburg. Hence the chosen time-frame enables analysis of publishing patterns before and after the introduction of the new research funding system.

Results: The empirical findings of this study suggest that for all types of publications (except reports) the number of publications in a five-year period after the introduction of PRFS (2010-2014) was greater than in the period before (2005-2009). More detailed analysis reveals that the number of peer-reviewed journal articles in particular has increased at a much greater rate after the introduction of PRFS.

The main conclusion from interpreting the findings in relation to the quality of educational research is two-fold: if publishing patterns are interpreted merely according to assumptions of the PRFS model, then a shift towards quantity instead of quality is foregrounded. In contrast, if specifics of educational research are included in the interpretation, then the publishing patterns may well indicate reduced quality of educational research.

Superficially the above interpretation may indicate compliance with assumptions implicit in PRFS. Yet the presence of other patterns in the data that contradict the PRFS assumptions suggests that change in publication patterns cannot be linked with the use of PRFS in any straightforward sense.

Acknowledgements

In writing this thesis, I have become entirely sure of one fact about research – research requires *collective* effort. I would not have arrived at this point without all the people who have in one or another way supported me in writing this thesis.

My first thanks goes to historian, philosopher, sociologist or, in other words, *thinker* Steve Fuller, whose writings have inspired, influenced and kept me intellectually 'irritated' through-out the work on this thesis. The playfully creative insights that I found in his 'Social Epistemology' (2002), 'Science' (2010), 'The Knowledge Book' (2007) and in his many other books, articles and book chapters have helped me to understand which ideas are those that I care about most and, more importantly, what other considerations may be involved in understanding such ideas.

In the same way, I am thankful to people who have contributed more directly in the route towards this thesis: my supervisors Sverker Lindblad and Gustaf Nelhans, whose critical comments were most helpful, most annoying and yet inspirational. Also, thanks goes to all the tutors who in their different ways how showed what does it mean to do research: Ilse Hakvoort, Dennis Beach, Girma Berhanu, Astrid Schubring, Anna Maria Hipkiss, Miguel Garcia Yeste. Discussions with all these people have been essential for the development of the ideas that I present in this text. Especially, I appreciate the lengthy and inspiring discussions with Kajsa Yang Hansen who had showed how exciting can be the process of tackling the endless conceptual and technical difficulties that may emerge when working with numbers. Similarly, the discussions in courses in Theory of Science led by Johan Söderberg have been a major step towards clarity in thinking on matters I care about.

Special thanks I owe to Dawn Sanders, Petra Angervall and Åke Ingerman whose stories on academic publications and their role in present academia were the starting points for this study. Then, in autumn 2014, I did not suspect that scholarly communication and bibliometric research methods may become my main academic interest.

Also, I greatly appreciate support and advice from the library of the University of Gothenburg: Bo Jarneving and Sofia Gullstrand, and from the Faculty of Education Administrative Director Ulrika Nordberg Petersson.

My family also deserves special thanks. This thesis would not have been possible without the endless encouragement, support and critical comments from the closest ones.

Contents

Abbreviations

HE	Higher education
GU	University of Gothenburg
GUP	University of Gothenburg database ‘Gothenburg University Publications’
PRFS	Performance-based research funding system
$PRFS_{GU}$	Performance-based research funding system that is used at the University of Gothenburg
SSH	Social sciences and humanities
WoS	Scholarly publication database ‘Web of Science’
SE	Social epistemology
STS	Science and Technology Studies, also known as Social Studies of Science

Figures

Tables

Appendices

Introduction

Since the early days of educational research in the early 20th century, this field of research has continuously been a home for a more or less harsh debate on what educational research should strive for and what would be the best ways to achieve the goals set (e.g. Lagemann, 2000; Baez & Boyles, 2009; Furlong & Lawn, 2011; Depaepe, 2002). In this debate, the quality of educational research has been one of the core topics, a topic in which epistemic and social concerns are blended. On the one hand, epistemic concerns have been expressed about the extent to which educational research might be deemed a trustworthy source of knowledge (e.g. Torgerson & Torgerson, 2001). On the other hand, concerns have been expressed about the social role that educational research may (and may not) play for teachers (e.g. Winch, Oancea, & Orchard, 2015) and policy-makers (e.g. Welch, 2015). In the more recent commentaries, more and more questions have been raised about this interplay of the social and epistemic, both of which are seen as inevitable aspects of the quality of educational research (see a compilation of in-depth discussion pieces on the topic by Reid, Hart, & Peters, 2014). These foundational aspects of educational research have become especially urgent in the present performance-oriented contexts of this field of inquiry.

Over the last two decades more and more countries have taken the performance-oriented-route in the higher education sector (HE henceforth) (e.g. Elzinga, 2012; Hicks, 2012; Ball, 2012). In present HE systems, more and more countries and institutions have introduced so-called performance-based research funding systems (henceforth PRFS) (Hicks, 2012; Pajić, 2015). In these systems, a portion of research funding is distributed according to quantitative indicators of performance, with bibliometric indicators being commonly used – those that are linked with scholarly publications. Examples of such indicators are counts of publications and citations.

In these systems, it is implied that funding flows to those who perform best (Hicks, 2012). Indeed, for example, in the Swedish Government proposition of the model for research funding distribution that was introduced in 2009, it is stated that the purpose of the system is to 'encourage higher education institutions to find research-profiles that give a competitive advantage in relation to others' (*Prop. 2008/09:50*, 2008, p.23, my translation). In this system, one of the indicators in which the HE institutions are encouraged to compete is number of publications and their citations.

This assumption of funding-flow privileging high performance forms the core problem with performance indicators. In this context of problematising PRFS assumptions, I should note that in this study I focus on bibliometric indicators only. Other performance indicators such as the number of research grants, the number of publication views or profile-page views in social networking platforms such as Academia.edu are not discussed here. For this reason, I will from now on refer to PRFS only with reference to bibliometrics-based research funding systems as a sub-type of the just introduced PRFS.

More and more critical insights are voiced about the high risk of abusing bibliometric indicators of performance. It is suggested that the use of PRFS may have potential implications to knowledge produced in such performance-oriented contexts (e.g. Rijcke, Wouters, Rushforth, Franssen, & Hammarfelt, 2015; Hammarfelt & de Rijcke, 2015; Aagaard, 2015). For example, in a context of PRFS, change can be observed in publishing patterns (Sivertsen, 2010; Moed, 2008; Butler, 2003). There is a tendency for those disciplines that are incompatible with PRFS to change their approach to publishing and by doing so – to change the actual knowledge practice.

This incompatibility with PRFS might precisely be the case for educational research. Let me explain. Incompatibility with PRFS is linked to publishing traditions that vary among academic disciplines (Sandström & Sandström, 2009). In educational research, as within most social sciences and humanities (henceforth SSH), the publishing traditions are very specific in terms of the most commonly used publication type and language (Sivertsen, 2016; Engels, Ossenblok, & Spruyt, 2012; Nederhof, 2006). One may find publications in a form of peer-reviewed journal articles, books and book chapters that may be written not only in English – the *lingua franca* of present academia – but

also in national languages. These patterns are exactly the case for educational research (Sivertsen & Larsen, 2012; Engels et al., 2012; Dees, 2008; Diem & Wolter, 2013; Hansen & Lindblad, 2010).

These patterns turn out to be a source of ambiguity on what is the quality of educational research in a context of PRFS. In performance-oriented contexts, great emphasis is put on peer-review journal articles. This is done for two reasons: either it is assumed that peer-review ensures a more or less reliable warrant of quality (De Bellis, 2009) or the data that are used in PRFS employ international databases of scholarly communication such as Web of Science (WoS) or Scopus, in which peer-reviewed journal article is the publication type with the highest coverage. Consequently, in a context of PRFS, priority is given to publishing peer-reviewed journal articles. While the role of this medium may be communication of knowledge, it may equally turn out be merely a currency that is used for enhancing performance scores and so explain the increasing numbers of peer-reviewed publications in the contexts of PRFS (Butler, 2003; Moed, 2008; Sivertsen & Larsen, 2012).

If I take these considerations back to educational research and the quality of this research, it may be expected that in the presence of PRFS, educational research publishing patterns would indeed change according to the priorities within PRFS, as has been the case in, for example, Flanders – the Dutch-speaking part of Belgium (Ossenblok, Engels, & Sivertsen, 2012). But even if such changes do occur, it is still not quite clear how change is then supposed to relate to the quality of educational research. It may therefore seem that publishing patterns and the quality of research are two different areas of consideration. Yet, I argue that in the context of PRFS, quality and output become so closely related that it is no longer clear what the actual interaction between the two is; or in other words, the key question becomes, *what implications to the quality of educational research may be identified in the changes of publishing patterns that occur in a context of PRFS?*

This is the aim that forms the point of departure for this theoretically-oriented study in which I aim to explore a particular case – the publishing patterns of educational research (2005-2014) at the University of Gothenburg (GU) in a context of the performance-based research funding model that is used to distribute the funding within GU ($PRFS_{GU}$ henceforth) among faculties. This period of ten years (2005-2014) is chosen purposefully: $PRFS_{GU}$ was implemented in 2009 (University of Gothenburg & Universitetslednings kansli, 2008, sec. 10). Therefore in this study, I will compare the publishing patterns in a five-year period before the introduction of the model (2005-2009) and after (2010-2014). The focus is on the exploration and interpretation of change in publishing patterns in relation to the quality of education research. How exactly I achieve that goal, I will explain in the chapters to come.

Now, it is time to introduce my own position and write a few notes on the significance of this study. My position in this study can be described as the so-called 'interested non-participant' position (Fuller, 2002). I care about education, and I believe that educational research is (or can be) a worthy source of knowledge about and for education, but at the same time I am aware of the political aspects of normative claims about knowledge, whereby one claim can benefit some at the expense of some others—knowledge is never neutral with respect to social difference. Therefore, my own intention is to better understand what may be taken as evidence for educational researchers doing their work well.

The significance of this study dwells in the distinctive nature of educational research. The quality of educational research is a highly problematic, yet central concept, but in the present performance-oriented context it becomes urgent to conceptualise changes in this field occurring under the influence of external criteria, given that these seem to directly affect how the worth of educational research is assessed (as in PRFS). It may be that the change is beneficial for educational research, but it may equally be that there is a risk of 'damaging the system with the very tools designed to improve it' (Hicks & Wouters, 2015). Hence, the theoretical insights from this study may be of value in discussions and scholarly reflections on the desirability of performance-driven shifts in educational research. At the same time the study's empirical evidence may separately stand as a contribution to those studies aiming to capture and understand shifts in publishing practices in a context of PRFS.

The Structure of the Thesis

The study just introduced is presented according to the following layout: I begin with the theoretical framework of the study – social epistemology (Fuller, 2002). I provide a rationale for each of the chapters at their beginning, but let me present a brief and more general guide to what follows in the rest of this thesis. The reader may have already noticed the length of this thesis and may wonder whether it is reasonable for a case study that explores bibliometric data. My reply is - yes (although, it is open for a discussion), this study does require the lengthy and relatively detailed tour through the philosophy of science, the history of bibliometrics and meta-debates in educational research that follows in the next chapters in order to *conceptualise* publishing patterns in a context of PRFS in relation to the quality of educational research. The focus of this study is on the theoretical understanding of the patterns explored empirically, hence the emphasis on theory.

In the next section, I turn to the literature review. In the final part of the literature review, I identify the gaps in the literature and describe the contribution of this study in a greater detail. After the literature review, I present the research questions that will guide this study. In the section following, I describe the research design and the method of this study. Next, I describe the context of this study - the University of Gothenburg. Then, I provide a detailed description of the $PRFS_{GU}$ that is used to calculate funding for the Faculty of Education: the so-called Norwegian model. After that, I report findings. The final chapter of the study contains discussion of the findings of this study. I begin with a discussion of the empirical findings – the publishing patterns. Then, I discuss these findings more particularly and theoretically in relation to the quality of educational research. In the final part of the discussion, I present the main conclusions of this study and briefly reflect upon some limitations of this study and the possible direction for further research.

Social Epistemology

In this section, I describe the theoretical framework of this study, which is derived from Steve Fuller's (2002) social epistemology (SE henceforth). The rationale behind this early introduction of my theoretical framework is conceptual. I reject the idea that meanings are possible without context and prior assumptions about what the world is like (ontology) and indeed in which we might come to know about the world (epistemology). In doing so, I distance myself in particular from naïve realism, such as the idea that an objective reading or assessment of scholarly texts is possible, which is why I regard it as crucial to introduce and describe here the way I think about the main research object of this study - the change in publishing patterns of educational research in a context of PRFS.

Social epistemology (Fuller, 2002) is an approach to study knowledge practices and it is closely related to a much broader field of inquiry commonly referred to as either Social Studies of Science or Science and Technology Studies (STS). A key assumption within SE and most of STS is that research is seen as social practice (e.g. Sismondo, 2010; Fuller, 2010). It is supposed that the ways in which knowledge is acquired can and should be studied in its own right, employing research methods commonplace in sociology, anthropology or, as in this study – in library and information science.

Steve Fuller's approach to the exploration of knowledge practices can be distinguished from closely related approaches – for example, the social epistemology by Alvin Goldman (e.g. 1999). The key distinction between Goldman's and Fuller's versions of social epistemology is that Goldman's social epistemology is embedded in a tradition of analytic philosophy, where knowledge is assumed to be an entirely individual phenomenon; as opposed to Fuller's view of knowledge being an entirely social phenomenon (Fuller, 2007, pp.1-5).

To make more clear how research may be considered a social practice, I will undertake a slight detour into history of STS and SE that leads us back 1960s and the work of science historian Thomas Kuhn and his concept of a research 'paradigm'.

On Paradigm

Paradigm is a concept that Thomas Kuhn introduced in his 'The Structure of Scientific Revolutions' (1970) and this concept is central in this study for two reasons. Firstly, it helps to contextualise Fuller's SE and, secondly, it adds to the vocabulary that I need to discuss the various ways of understanding the quality of educational research – a topic that I discuss in the next chapter (page 10).

Kuhn's 'The Structure of Scientific Revolutions' marks a breaking point in the understanding of science and in a way sets a foundation stone for STS. Kuhn conducts a historiography of events during what is assumed to be 'the Scientific Revolution' of the 17th century. He shows that much of what appears to us as linear progress in scientific thought is misleading and derived from a misconception of scientific effort. Introducing the terms 'paradigm' and 'normal science', Kuhn suggests that in the pursuit of knowledge, researchers are most of the time engaged in a sort of puzzle-solving activity, an activity in which main disciplinary doctrines on ontology and epistemology are settled. In educational research, this could be exemplified with research on intelligence tests (e.g. Valentin Kvist & Gustafsson, 2008). If it assumed that there is such a phenomenon as intelligence and that it is possible to acquire new knowledge about intelligence through tests, then the further task for research is pretty clear – more tests in more contexts and with more sophisticated methods have to be conducted in order to know more about intelligence. Such a progressive line of inquiry is what following Kuhn would be called 'normal science'. Within normal science, new knowledge is acquired under shared ontological and epistemological assumptions, which consequentially determine the choice of problem areas, research questions and research methods. All these assumptions taken together constitute what Kuhn refers to as a 'paradigm'. It is therefore not the methods that are used within a particular problem area that determine a paradigm. Instead, for Kuhn a paradigm is a whole worldview, containing both ontological and epistemological assumptions that guides understanding within a given field of what research is and what the field collectively strives to achieve.

One may argue that Kuhn's approach to theoretical questions on the basis of inquiry into actual research is still the essence of STS. Within STS, the value of the exploration and theorisation of the actual research practice is what is referred to by the assumption that research is a social practice. I will explain this in greater detail shortly, but first I return to SE, the theoretical framework of this study.

Most of the premises of SE, I have acquired from Steve Fuller's book entitled 'Social Epistemology' (2002) – this is the main book in which he elaborates his approach to study knowledge practices. He himself describes his approach as:

> *...a naturalistic approach to the normative questions surrounding the organization of knowledge processes and products. It seeks to provide guidance on how and what we should know on the basis of how and what we actually know.*
>
> (Fuller, 2007, p.177)

The two key features of Fuller's SE that need further explanation are the notions of the 'naturalistic' and the 'normative'. The 'naturalistic' refers to 'the idea that knowledge cannot be about the world unless it is clearly situated in the world' (Fuller, 2007, p.108). With such a claim Fuller acknowledges that knowledge practices are knowledge practices *in relation to* other social practices. It might be easier to understand this idea via the more straightforward assumption that knowledge is social.

When I say that 'knowledge is social' I mean that knowledge becomes knowledge only in relation to others. The fact that I know *p* is dependent on a community that, firstly, recognises *p*, and then secondly, recognises the distinction between knowing and not-knowing *p*. Without such an aware community my knowing *p* is meaningless.

Fuller in another text develops this idea by saying that knowledge is a positional good. Knowledge 'is supposed to expand the knower's possibilities for action by contracting the possible actions of others. These 'others' may range from fellow knowers to non-knowing natural and artificial entities' (Fuller, 2003, p.107). He continues that 'differences arise over the normative spin given to the slogan: should the stress be placed on the *opening* or the *closing* of possibilities for action? If the former, then the range of knowers is likely to be restricted; if the latter, then the range is likely to be extended' (Ibid.).

Here becomes apparent the *political* nature of knowledge and knowledge practices. Fuller argues that 'the social acceptance of a knowledge claim always serves to benefit certain interest groups in the society and to disadvantage others' (Fuller, 2002, p.10). He therefore also suggests that 'granting epistemic warrant is a covert form of distributing power' (Fuller, 2002, p.10). This thought is central for this study, since it illuminates the political nature of a discussion about the quality of educational research.

The just introduced points have several consequences for the ways one may think about the implications of changes in publishing patterns that occur in a context of PRFS for the quality of educational research. The first, and the most important, consequence is that the 'quality of educational research' should itself also be considered an inherently social concept.

The idea that knowledge about the quality of educational research is inevitably social might seem ambiguous and controversial. The concept of quality has always been highly problematic and complex and particularly so in the field of educational research (Pring, 2015). There have been attempts to define quality in relation its epistemology – the theoretical grounds; that is, grounds that are taken to be internal to science and scholarship itself— that would render a particular claim into a knowledge claim. This is what I call the *epistemic* aspect of quality. In the same way, there have been attempts to argue for research quality in relation its originality, significance, social value and relevance. Such arguments refer to attributes that are taken to be external to science and scholarship, and so may be more typically understood as the *social* aspect of quality.

In the framework of social epistemology however, such a distinction between epistemic and social is false. Both of these are considered social for reasons that can be traced back to ideas of philosophers such as John Locke (Locke, 1990[1689]). According to Locke, the epistemic is prior to the social.

Hence Locke's view opposes what is argued by Steve Fuller in social epistemology, and therefore it is crucial to understand how the quality of educational research is seen in this study. To show what I mean, I will briefly note how Locke's legacy has evolved and how it is linked to SE. I will not take this long journey from Locke's primary/secondary qualities to Fuller's SE, but I will instead present the key arguments that are crucial for this study.

Locke in his 'An Essay Concerning Human Understanding' made a distinction between primary and secondary qualities of 'bodies'. Locke (1990) suggests that primary qualities are those that 'are utterly inseparable from the body, in what state soever it be' (p.89). In contrast, secondary qualities are those 'qualities which in truth are nothing in the objects themselves but power to produce various sensations in us by their primary qualities' (Ibid. p.90). If transferred to the context of this study, such a distinction might suggest the following. The quality of educational research in the sense of primary quality may be identified by some intrinsic and context-independent constituents of research or what I introduced as the epistemic aspect of quality. In contrast, in the sense of secondary qualities, it would refer to the social aspect of quality – the characteristics that 'produce various sensations'. Hence according to Locke, the epistemic is prior to the social.

Locke's legacy in the later centuries received critique and was greatly transformed (e.g. Kant, 2004[1783]; Hegel, 2013[1807]). First, Kant (2004) showed that Locke's distinction between primary and secondary qualities is false. He argued that both primary and secondary qualities are categories developed in mind. Without mind there would be no possible experience of whatever a concept may refer to. Yet, for the concepts, there may be no resemblance to the world whatsoever, because knowledge can be acquired only through categories within mind. Hence Kant in turn rejected the very possibility of knowledge about the world as it is.

Kant's objections were taken further by Hegel. Such a sceptical position did not satisfy Hegel (2013), who proposed that if Kant is right, then the very fact that one may agree with such a sceptical position towards knowledge, already suggests that people have an in-built sense of movement towards truth, but this movement is dialectical. This means that the so-called epistemic criteria of knowledge may be adjusted as knowledge is acquired. At the point when epistemic criteria are adjusted, those criteria that served as the basis for knowledge become merely social criteria from a perspective of the present criteria. Therefore, according to Hegel, if I say that the social and epistemic are both social, I do not promote the status of social criteria to epistemology. Instead, I acknowledge that the process of acquiring knowledge about knowledge is dialectical and hence epistemic criteria are social.

A direct link can be drawn from Hegel to Fuller's social epistemology and his assumption that the knowledge about the goals of knowledge is acquired dialectically. It means that more knowledge on a particular topic can lead to readjustments of goals of knowledge practice. An example could be the emergence of the feminist research paradigm in the 1960s. In this paradigm, the emphasis is set on the emancipation and empowerment of women (Punch, 2005, pp.136-8). Such understanding of a feminist knowledge practice goal might however not have been possible in, say, the late 19th century – the early days of educational research. A lot more was yet to be known before feminist research was possible (see Mertens, 2010, pp.15-21).

From the above follows that, within the conceptual scheme of social epistemology, definitions of quality can only be founded on the social contexts in which they are embedded. The social epistemologist supposes that the meaning of the concept of quality is deeply entrenched in the context in which it is used and cannot somehow be divorced from it scientifically—for example by foregrounding a science-internal view of its purported 'purely' epistemic aspects, such as would be imagined in philosophy—because such an attempt is itself also a particular social practice embedded in context. This is where the notion 'naturalistic' becomes central. The naturalistic aspect of social epistemology then means that in order to understand the knowledge practice that takes place, one may consider knowledge of the actual conduct of this practice. In this way, Fuller distances himself from analytic philosophy of science where the principles guiding research are theorising without reference to the actual possibilities to implement such principles. In less abstract terms, it means the following.

In this study, I explore two contexts. First, I explore the internal context of educational research that refers to the diversity of ways how the quality of educational research may be understood. Second, I explore the external context of educational research – the use of PRFS in HE. On the one hand, one may identify concerns that could be rendered philosophical – what is educational research, what its aims should be and what may be the ways of achieving them. On the other hand, the existing knowledge on which goals have led to what results is what Fuller (2002) proposes as worthy sources of understanding whether the present conduct of educational research is acceptable. Such concerns are what Fuller indentifies as the naturalistic aspect of social epistemology – the acknowledgement of the value of empirical accounts of the actual conduct of knowledge practices.

Now, I proceed to the description of Fuller's notion of normativity - the second key feature of SE, which partially I have already introduced above. The normative part of social epistemology can be characterised with a principle from past to future (Fuller, 2002, p.24). It means that 'normative judgements . . . about past are meant as the basis for issuing normative judgments . . . about future knowledge production' (Fuller, 2002, p.24). In this way, the description of knowledge practices may be approached without presuming the goals of the practice in question.

Such an idea is based on two complementing theoretical assumptions: firstly, that knowledge can be acquired deductively and secondly, that the acquisition of knowledge about the goals of knowledge practices is a dialectical process. The dialectical process was introduced above, but now I explain what is meant by deductive approach to knowledge acquisition.

Fuller proposes that knowledge about the goals of knowledge can be acquired in the following way: a proposition of a knowledge claim is made and then follows an attempt to refute it. In this way Fuller's approach in his social epistemology draws closely to philosopher of science Karl Popper, who in the mid 20th century proposed falsifiability as a general principle in theory of knowledge (Popper, 2002). Popper sought to address the problem that there are no means to be sure that new knowledge is closer to truth than previous knowledge, if the knowledge in question is acquired through induction – from individual observations to generalisations. Popper proposed that, if a proposition for knowledge claim is created in a way that enables its falsification, then the knowledge-making avoids the trap of continuously reinforcing false beliefs and assumptions, as may occur when research is guided instead by the principle of verification.

In the context of this study, it means that if I wish to explore goals of the present educational research I can do so by stating a hypothesis in a form 'the goal of educational research is x' and then identifying arguments for why such goals may not be achievable. In this study, my interest is in educational research is conducted in a context of PRFS. Therefore, I focus on goals for educational research that can be achieved in this particular context. I will return to this shortly.

On Doing Social Epistemology

The question that still remains is how social epistemology is done. In brief, pursuit of social epistemology is done by 'evaluating the metaphysical scheme of the individuals that [her] own metaphysical scheme classifies' (Fuller, 2002, p.35-6). It means that I employ my conceptual sources (my metaphysical scheme) first, to make sense of knowledge practice that I am interested in. In doing so, I classify it. I create an interpretation of the metaphysical scheme of this knowledge practice. Having done that, I evaluate this (interpretation of the) metaphysical scheme. How can I recognise these 'metaphysical schemes'? Fuller explains:

> *[t]he basic concepts and principles of social epistemology are developed and justified in the actual contexts of knowledge production that concern the social epistemologist. Thus, one starts in medias res, treating current knowledge production practices as empirical constraints on the possible directions that future knowledge production can take.*
>
> (Fuller, 2002, p.xv-vi)

It means that I refer to the two contexts I am interested in – the internal and the external contexts of educational research. Yet, I remind, for the external context I am interested only in the use of PRFS.

Further, Fuller describes the sources of these metaphysical schemes as linguistic and symbolic practices in which a particular group of individuals are engaged in 'collective representation' (Fuller, 2002, p.54). This collective representation 'arises not when everyone has the same beliefs, nor even when everyone believes that a belief has been accepted by the group; rather, it arises when everyone tacitly agrees to express *whatever they may happen to believe* in terms of specific linguistic and other symbolic practices' (Ibid).

Translating such reasoning to the context of my study, I see publishing practices within educational research as a part of the collective representation of this research field. There are various meanings that can be attached to publications (Schaffner, 1994), but in this study publications are treated as codified practices whose meaning is context-dependent and the given context is educational research conducted before and during the PRFS scheme of research funding. In this way I distance myself from a view that it is plausible to define a publication in transcendental terms that span various temporal and spatial contexts. Fuller suggests that books seen in this way are merely a 'language game one must play in order to pass off an interpretation of the text as legitimate to a given community of readers' (Fuller, 2002, p.53). Following Fuller's thoughts, publishing practice is a codified practice in which researchers act in relation to publications, in whatever way these may be defined. By 'act' I mean read, write, publish, cite and in any other way meaningfully engage with publications.

One may wonder, if there is no meaning inherent in the idea a publication in itself, then how can publications be identified? The answer to such a question is, through language. If the word 'publication' is used in a sentence, then I assume that this particular sentence (and maybe be other to which this one linked) is related to publications.

A further crucial point is that publishing remains a codified practice also beyond the context of university research. This is one of the consequences of treating knowledge practices as social practices. The internal and external contexts of the educational research that I have described can exist because there is a context in which research is recognised as research and actions take place accordingly. Returning to publishing practices, a book is not a book only for its authors and, for example, bibliometricians within universities. It is a book also for its readers which may be found beyond university walls. The same book is a book also for teachers, students, policy-makers or any other group that may be interest in a particular text, and it is still a book for all those who are not interested in it. But it might not be hard to imagine that the ways a book is valued by authors, teachers or any of the other social actors are not the same. What could be called as a network of interacting symbolic practices is a crucial point to keep in mind, when conceptualising the implications (or the costs and benefits in Fuller's terms) from change in publishing patterns. The interpretation of change that occurs within a university by scholars may be independent from, for example, the interpretation of this change by teachers. These diverse and yet overlapping metaphysical schemes and the potential tension among them is exactly what constitutes the categories that I will work with in this study.

It may be time to remind the reader that in this study I am interested in the use of PRFS as a specific context for educational research. Thus I conceptualise publications, research and its quality in relation to the concepts used within PRFS. PRFS is the concept that constitutes the external context of educational research. In order to be able to link PRFS with publications and with the quality of educational research, few more notes are needed on PRFS and its dialectical nature. What follows, is my adaptation of Fuller's SE to the context of this study.

In my reading, PRFS are part of two separate codified practices. One is the constitutive and the other the regulative practice of research. The regulative aspect of PRFS refers to the arguments used to justify the use of such a system. Typically, these arguments are found in research policy (Hicks, 2012). It is these arguments, in which I identify the assumed goals of knowledge practices to which PRFS are applied.

Among the more common arguments that are named in research policy are the enhancement of accountability and the competitiveness of universities (or any other units in which a PRFS is used) in the knowledge economy (Hicks, 2012). The first is achieved by the introduction of continuous monitoring of research systems by means of bibliometric or other performance indicators. The second – increased competitiveness – is achieved by introducing a reinforcement mechanism (funding, status) that rewards particular kind of knowledge practice. This second argument of increased competitiveness is where it is possible to identify the implicit assumption of the goals of knowledge practices that feeds the constitutive part of PRFS.

The constitutive aspect refers to the assumptions that are implicit in PRFS, with the reinforcement mechanisms identified above reifying a particular notion of research quality. In this sense, PRFS plays a role in constituting a knowledge practice in a particular way. Sandström and Sandström (2009) named this aspect the 'incentive system' (p.248).

Although in research policy universities are not always explicitly asked to adjust their knowledge practices to fit the performance indicators, universities are explicitly asked to enhance the quality of research (for Sweden see e.g. *Prop. 2008/09:50*, 2008), but recognition that the concept 'quality' is highly complex and may greatly vary in various national, institutional and disciplinary contexts, is hindered by the reification of a particular reductive projection of research quality in PRFS. Hence, the regulative aspect of PRFS acquires characteristics that may render PRFS constitutive. In a way, research quality is replaced with a set of indicators that may or may not coincide with one or more established senses of quality present in a particular context. Even where in research policy there is no explicit goal to increase the score in particular performance criteria, implicitly such performance scores still become the goal, since in PRFS the only evidence permissible for showing 'enhanced quality' are performance indicators. Hence even if bibliometric performance is not explicitly equalled with quality, the implicit invitation in PRFS is to prioritise certain activities, such as writing and publishing peer-reviewed articles. Thus, actions that may have been contributing to the quality of educational research, but contradict the actions prioritised in PRFS, are potentially hindered.

This is the entrance point in this study for the implications of the use of PRFS to the quality of educational research. Fuller suggests that the main question that guides the social epistemologist in her exploration of knowledge practices is: 'What sorts of goals can be realized given the actual structural constraints on knowledge production?' (Fuller, 2002, p.27). Accordingly, my focus is the goals that may be achieved given the conceptual scheme implicit in PRFS acting as structural constraint on educational research. Then one may ask 'whether we wish the future to continue certain tendencies of the past, and, if so, which ones?' (Fuller, 2002, p.xv-vi). This is how the normative is achieved within social epistemology. First, knowledge practice is described, the possible goals that can be achieved in such practice are identified, and then there may be a discussion on whether having such goals is desirable. How such considerations manifest in the case of GU, is an open question that I explore in the rest of this thesis. In the next chapter, I proceed with an exploration of scholarly literature.

Bibliometrics and the Quality of Educational Research

This chapter serves two purposes. On the one hand, I review and draw together insights from scholarly literature with a goal to show the diversity within the bibliometric studies of educational research and to identify what is known about change in publishing patterns in a context of PRFS, and what may be the relationship between such changes and the quality of educational research. Each of these themes I present in a separate section. In this review, I identify gaps that allow me to formulate the aim of this study in a form of research questions. Research questions are presented on page 27.

On the other hand, this chapter serves also as a target for the theoretical framework of this study. In the previous chapter, I introduced the main theoretical attributes pertaining to the quality of educational research and changes in publishing patterns in a context of PRFS. I noted that '[t]he basic concepts and principles of social epistemology are developed and justified in the actual contexts of knowledge production' (Fuller, 2002, p.xv-vi). In this study, my interest lies in two such actual contexts – the internal and the external context of educational research. Hence, in this chapter I identify the conceptual categories constituting each of the two contexts: in relation to the external context, I explore what concepts are linked to publications, the use of PRFS and the quality of educational research. With respect to the internal context, I focus on the quality of educational research and my selection of arguments is guided by considerations of the various ways in which quality is conceptualised in this field of inquiry.

The literature on bibliometrics I have found in library and information science, social studies of science and sociology of higher education. Since the bibliometric literature is substantial and spread across disciplines and cannot be easily reviewed exhaustively or systematically, I have focused on three themes: (1) empirical exploration of publishing practices in educational research, (2) empirical exploration of change in publishing practices in a context of PRFS and (3) conceptualisation of such changes in relation to quality. I have identified key reading by key word searches ('quality', 'relevance', 'educational research', 'social epistemology') via Google Scholar, Web of Science, GU library search engine. Some further texts I have identified by browsing through journals or in the references of the texts I have been reading, while still others are articles that were recommended to me.

The literature on the quality of educational research I have identified in much the same way. Special attention is paid to the literature in the philosophy of education and sociology of education and higher education. In this way I have also stayed close to Steve Fuller (2002), who himself draws mainly from literature in sociology and philosophy. This particular disciplinary combination enables a discussion about the goals of research (philosophy) and the constraints to such goals in practice (sociology).

In the discussion of each of the topics, I follow the same scheme: I begin with general arguments on the topic and then I turn to Sweden and GU as a particular case. The first section in the review of the literature is on bibliometrics. I begin with a brief introduction in the origins of bibliometrics as a research method. I focus here on studies that have explored publishing patterns within the field of educational research. I continue on the use of bibliometrics in PRFS. Then, a section on the quality of educational research follows. In the final part of the chapter, I draw together insights from the previous sections to identify and explicate the role of my study in relation to the current state of knowledge on the topic.

The Tale of Bibliometrics

Unobtrusive Bibliometrics

I begin with a brief historical detour on the origins of *bibliometrics as a research tool* that has shed some light on publishing patterns of educational research. The beginnings of bibliometrics are typically linked to work done in 1960s US by Derek de Solla Price and Eugene Garfield. Price did a series of studies exploring the escalating numbers of publications and attempted to capture the

patterns employing mathematical equations (e.g. Price, 1971), while Eugene Garfield created the science citation index – the same that is part of the most popular and widely used academic publications database Web of Science (WoS). A characteristic of the early work within bibliometrics was that bibliometric data were treated as 'relatively unobtrusive sociometric data' enabling quantitative exploration of processes within science (De Bellis, 2009, p.50, emphasis added).

This assumption of bibliometric analysis being an unobtrusive approach to study processes of science forms the central point of widespread discussion of the negative implications – also called the *unintended consequences* – that spring from the use of PRFS (e.g. van Dalen & Henkens, 2012; Bornmann, 2011; Laudel & Gläser, 2006). I will return to this topic shortly, but first a few more words need to be said on bibliometrics being unobtrusive.

The particular understanding of bibliometrics being an unobtrusive research method can be traced back to library and information science. In this discipline, various methods have been developed to explore and describe various aspects of publishing patterns or processes in science without interfering into the actual processes. The range of topics explored under such assumption is wide: from simple descriptive summaries of bibliographic data of disciplines (Fernández-Cano & Bueno, 1999), countries (Moed, 2008), topics to various statistical techniques to explore co-authorship patterns (Bebeau & Monson, 2011), internationalisation (Verleysen & Engels, 2014) and many other topics.

Studies that have explored publishing patterns of educational research suggest a rather great diversity in various national contexts. A study from Norway explored these matters employing their national bibliometric database that covers the total volume of the Norwegian academic publications (2005-2009) and WoS (Sivertsen & Larsen, 2012). Their findings suggest that peer-reviewed journal articles in educational research constitute only about a half (46 per cent) of the total volume (N=2396). The next type – book chapters – is the most popular type within Norwegian educational research, accounting for 49 per cent. The final type of publications that was explored in the study is books. In the educational research field, 9 per cent of the publications are books. Thus, it seems indeed to be the case, that, at least for Norwegian educational research, the peer-reviewed journal articles is not the dominant publishing channel – it accounts for roughly about a half of the total number of publications for a period of 5 years. In addition to these patterns, the study also showed that patterns within educational research differ greatly not only from those in natural sciences and medicine, but also from the patterns in most SSH.

The just cited study from Norway has employed national database that covers all publications, but most commonly, bibliometric studies employ commercial international databases of scholarly publications such as WoS or Scopus. In these databases, the main type of publications is peer-reviewed journal article. This can be explained by the fact that historically these databases were constructed for natural sciences where peer-review journal articles are the dominant publishing channel. Within educational research field as well as other SSH where the publishing is much more diverse, only a small part of publications is covered by these databases (Nederhof, 2006; Sivertsen & Larsen, 2012). Hence the possibility explore publishing patterns in educational research is dependent on accessibility to bibliometric data.

A comparison of the data from the national database and WoS in the study by Sivertsen and Larsen (2012) showed that the coverage of the educational research publication in WoS is rather surprising: only 9 per cent of the total volume within this research field is covered by WoS. The average for social sciences and humanities is 20 and 11 per cent respectively.

In the literature, there have been suggestions that databases such as WoS and Scopus are biased towards publishing channels (series and publishers) that favour English language (Nederhof, 2006). Fields such as educational research where emphasis may be on publishing for national audience may turn out in a marginalised position. Indeed, the above-cited study from Norway suggested that only 33 per cent of the publications within educational research were published in English, while the rest – in the national language (Sivertsen & Larsen, 2012). This pattern is not characteristic for all SSH. In these fields in general, the average share of publications in English is about a half of the total volume.

In a more recent study, Sivertsen (2016) advanced an argument that the increase in the number of publications in English – in other words, internationalisation – within SSH that is typically linked to publishing channel coverage within WoS or other databases can rather be explained as a return to the history of academic traditions. He suggests that publishing in SSH *originally* has been international – texts have been written in Latin and only later with the introduction of mass education in 20th century there has been move to publishing in national languages. Hence a common critique of PRFS as of pushing SSH towards international audiences might not be warranted. Although, he admits that 'the SSH would lose their ***raison d'être*** by disconnecting from the surrounding culture and society and by mainly communicating in international journals that are only read by peers abroad' (Sivertsen, 2016, p.359).

On the more recent communication patterns in SSH, Sivertsen (2016) reflected upon findings of his earlier study (Ossenblok et al., 2012) and concluded that 'publication patterns differ between the disciplines of the SSH while they are similar across countries within the disciplines' (p.359). Hence one may expect that publishing patterns of educational research would be similar in various countries. He continued that an investigation of change over time in publishing patterns in relation the publication type showed that the various publication types *do not* compete with others – peer-reviewed journal articles do not replace books or book chapters, but rather complement these more traditional types. Similarly, the analysis of publishing patterns on individual level suggested that the choice of language is not a matter of competition – most of researchers were bilingual in their publishing practices and none are identified as publishing only in national language. Finally, Sivertsen suggested that the publishing patterns in SSH are relatively stable over time.

A somewhat different suggestion could be identified in a comparative study that explored change over time (2005-2009) in internationalisation patterns in SSH in Flanders – the Dutch-speaking region in Belgium - and Norway (Ossenblok et al., 2012). As in Norway, in Flanders there is a regional database that contains all the publications of Flemish universities. These databases were the sources of the study. The study included data on publications in educational research, but it focused only on peer-reviewed journal articles (in Flanders n=369.1; in Norway n=1094). The study employed fractioned counting whereby a publication is divided by the number of authors, with 1/10 as the minimum fraction. From these publications, about a half of Norwegian publications (52.8 per cent) were published in national language, but for Flanders this number was only 15.5 per cent. In relation to the coverage of publications from educational research by WoS, the study reported that the share of WoS-covered publications has increased from 14.9 per cent (n=9.2) to 52.1 per cent (n=45) in Flanders and from 17.1 (n=29.4) to 19 per cent (n=52.6) in Norway. A comparison of the findings for Norway and the findings from the earlier discussed study by Sivertsen and Larsen (2012), which explored not only peer-reviewed journal articles but also books and book chapters, seems to suggest that the share of peer-reviewed journal articles in WoS is higher than the average coverage of articles together with books and book chapters.

Another study from Flanders (Engels et al., 2012) identified 1257 publications in Educational Sciences (as it is called in the original article) over a period of ten years (2000-2009). Of these, 92 per cent are peer-reviewed journal articles. Book chapters accounted only for 5 per cent, but books and edited books accounted for 0.05 and 0.06 per cent respectively, while proceedings constituted 2 per cent. Surprisingly, about 80 per cent of all these texts were written in English. Ossenblok et al. (2012) suggested that this might be related to the fact that prior to 2010 research funding in Flanders was distributed using only WoS data, thus prioritising publications in English.

Diem and Wolter (2013) conducted a bibliometric exploration of education sciences in Switzerland. The study explored scholarly publications in the period of 5 years (2005-2010). The inclusion criteria were linked to authors: the study included all the educational researchers that were employed in a Swiss university in 2010 (n=51). The study employed data from WoS and another popular academic bibliography database Google Scholar. A search through WoS identified 374 publications, of which 58 per cent are peer-reviewed journal articles. The rest were book reviews, editorials and other types which were not further described. In Google Scholar, the number of publications was much higher – 1159. Further details about types and language of these data were not provided. However, the findings

of this study suggested that despite the fact that Google Scholar covered four times more publications than WoS, it was not the case that Google Scholar covers all the publications that were identified in WoS. Only two thirds of the publications identified in WoS were covered by Google Scholar. In conclusion, the main drawback of the study is that it employed data from these highly problematic databases: it is not known how much of all Swiss educational research publications are covered by WoS data. In relation to Google Scholar, the data from this database are regarded of poor quality – the data are with errors and include as "publications" various sources that are regarded either as *grey literature* or the bibliographic information contains errors that can result in misleading results (e.g. Meho & Yang, 2007). Therefore, it is possible that the findings from the study by Diem and Wolter (2013) reflect only a minor part of the actual educational research publishing patterns within Switzerland and are misleading.

A study from Germany (Dees, 2008) reported that about a half of the total number of publications (n=4694) within German educational research (2004-2006) conducted at 15 institutions were book chapters (46.7 per cent). Journal articles accounted for about a third (33.4 per cent), but the share of books was 14.8 per cent. In this dataset, 88.1 per cent of all publications were written in German. However, authors suggested that the share of English-language publications varied among institutions: there were also one institution were about a half of all publications were in English. The data used in the study were acquired from the included institutions.

The variation in search methods and findings that was found in these and other studies indicate that the claim of publishing patterns within disciplines but across national contexts being more or less the same, as has been suggested by Sivertsen (2016), might not be conclusive and would require further exploration in more countries. The publication patterns in educational research seem to reflect substantial difference among countries, but this variation might be missed for the same reason that educational research is not widely explored by studies employing bibliometric methods: the relatively small field of educational research is not generally of interest.

For Sweden, a study by Hansen and Lindblad (2010) explored publishing patterns within specifically the field of educational research and it provided an approximate sense of patters of scholarly publication in this country. The study was conducted on a commission by the Committee for Educational Sciences (CES) within the Swedish Research Council. Within the study, the time frame for publications was 2004-2008 and further inclusion criteria were linked to authors, similarly to the study in Switzerland. The Swedish study included only those authors (n=650) who, firstly, from 2001-2007 had submitted at least one research grant proposal to CES and, secondly, who were employed by the University of Gothenburg, the University of Linköping or the University of Umeå. The number of grant applications from these three universities accounts for about a third of the total number, and the 650 researchers account for 27 per cent of the total number of researchers who submitted a grant application either as a main or participating application. Information on how these numbers relate to the total Swedish community of educational researchers that includes also those who do not submit grant applications is not provided. The publication data were obtained from university databases, the quality of which is not known.

About 4000 publications were identified for the period of 5 years (2004-2008). 23 per cent of these publications were peer-reviewed journal articles, book chapters – 25 percent, books – 5 per cent, edited books – 2 per cent, conference contributions – about 20 per cent, reports – 8 per cent, but doctoral theses – 2 per cent. A relatively large number (10 per cent) are publications categorised as 'Other article'. It is explained that a large part of these are 'reviews that are not aimed at the research community' (Hansen & Lindblad, 2010, p.36, my translation).

In relation to language, the study suggested that about half of the publications were in English, but Swedish was used for 44 per cent of the publications. The rest of publications were published in other languages. Peer-reviewed journal articles were mostly (88 per cent) published in English. The Swedish language was mostly used for edited books, reports and other articles. However, data on the language of publications were provided only for publications from Gothenburg and Umeå. The

authors indicated that the data on the language of publications from Linköping were not registered (Hansen & Lindblad, 2010).

A basic summary of publishing patterns at the University of Gothenburg can be most easily accessed via the University of Gothenburg database 'Gothenburg University Publications' (GUP; gup.ub.gu.se) which is open to public access. Here, a statistical summary that is expressed as full counts of publications of different type can be accessed with few *clicks* (see http://gup.ub.gu.se/statistics/). Even so, questions arise. What exactly is known about educational research at this university, when GUP shows that over the period of 10 years (2005-2014) 5030 publications are affiliated with the Faculty of Education at the University of Gothenburg? How to interpret the fact that of these 5030 publications, about one fifth (19 per cent; n= 971) are peer-reviewed journal articles, slightly less (17 per cent; n=847) – book chapters, but monographs – 3 per cent (n=147)?

It may be said that the patterns within GU are not identical to those in reported in the Swedish study by Hansen and Lindblad (2010). In GU, there seems to be a smaller share of peer-reviewed journal articles than in the average in the three Swedish universities over the period of 5 years (2004-2008). But in what sense, for example, are these periods and institutions comparable? In short, what do the numbers and variation in the numbers tell us about educational research and how it develops?

In relation to the findings from other countries, such comparisons seem even less meaningful – the categories of the publication types are not the same, thus any comparison of the numbers of share may be misleading. Thus, a more detailed study would be required that enables a richer contextualisation of such *raw* GUP data.

In summary, this review shows that publication patterns of educational research vary across studies and with the selection of data, variables and time-frames. It might seem that publishing patterns vary also across countries, yet the great methodological diversity in the reviewed studies does not allow straight-forward cross-country comparison. At the theoretical level that was introduced in the previous chapter it also seems reasonable to assume that publishing patterns will vary with the academic culture of the field *in a particular* context. This further increases the difficulty of drawing reliable conclusions from purely bibliometric data and studies *across* the various national and institutional contexts, and surely even more across disciplinary contexts – even among other SSH disciplines that are assumed to be similar in their publishing traditions. Keeping these insights in mind, I now proceed to my next theme: the use of bibliometrics in PRFS.

Bibliometrics in Performance Systems

The use of bibliometrics in PRFS characterises a great number of contexts in which educational research (and its publishing) takes place. In this section of the literature review, my intention is to identify what these systems are and how in these systems publishing patterns are linked to research quality or some related concept.

Earlier in the text, I referred to the use of performance-based research funding systems (PRFS) in HE as an example of the so-called performance-oriented turn. To understand the essence of this turn, one may recall how PRFS have come about and what reasoning supports the use of PRFS in HE.

Hicks (2012) refers to PRFS as research policy instruments. The understanding of these instruments is linked to the understanding of research policy: the targets set by policy, the challenges addressed by policy, and most importantly the local contexts of each individual PRFS. Drawing on a set of criteria proposed by Hicks, PRFS is a research funding system where funding depends on *ex post* evaluation of research on basis of research output. By *ex post* is meant evaluation after the research is conducted. What is meant by research output depends on each individual system, including bibliometric indicators such as number of publications and/or citations, or other more recently developed indicators used to measure research output. These other kinds of indicators are not discussed here.

The story of the use of bibliometric indicators for research funding distribution calculation can be traced back to the 1970s. At that time, bibliometrics as a method for research evaluation (or simply,

evaluative bibliometrics) were thought to be a more reliable replacement of peer review (De Bellis, 2009), which had been the traditional approach to research evaluation since the very early origins of modern science in the 17th century. It has been suggested that the rise of evaluative bibliometrics were due to growing doubts about scholars' ability to step over personal interests (De Bellis, 2009). As a quantitative approach to research evaluation, bibliometrics were seen as a promise for objective and impartial evaluation of research quality.

Such a suggestion is slightly controversial, since those publications that are most often explored using bibliometrics have gone through peer-review. Hence, doubts about the reliability of peer-review already imply doubting the reliability of evaluative bibliometrics. This controversy is linked to the various ways peer-review is used: the various 'levels' on which research is evaluated. On individual level, a typical example of a peer-review is evaluation of an article that is submitted to a journal. Here, a single piece of text is reviewed by one or two scholars. In the description of the WoS journal selection procedure, it is argued that peer-review 'signifies the overall quality and integrity of the research presented and the completeness of bibliographic elements, especially cited references' (Testa, 2016). Such a description corresponds to a claim that is common among bibliometricians, namely that research quality is such a complex notion that it can only be judged by peers (e.g. Van Raan, 1996; Schneider, 2009, p.367).

In contrast, when peer-review is applied on institutional, disciplinary or national level, then the argumentation for and against peer-review shifts from validity and reliability of this method to *costs.* Borrowing ideas from Theodore Porter (1994), a choice in favour of evaluative bibliometrics might be called a matter of *convenience.* It is much cheaper and easier to employ bibliometrics than the very costly peer-review. This has been the argument for introducing bibliometric indicators on national level in UK (Bridges, 2009).

It might therefore seem plausible to differentiate between individual-level and aggregate-level peer-review, but such a distinction might not acquire the same meaning in fields of research where there is great diversity especially in terms of paradigms. If peer-review either on individual or aggregate level would be applied within such a context, its meaningfulness would depend on the possibility of reaching consensus on research quality. Consensus is one of the key assumptions about what research is within the bibliometrics used within PRFS. I will return to consensus in greater detail in the discussion of findings of this study (page 47). Now, I proceed to an exploration of the actual methods that are used to evaluate research using bibliometric indicators, the conceptualisation of these methods, and finally how they have been transferred into performance-based research funding systems.

The two basic indicators that are used within evaluative bibliometrics are the number of publications and the number of citations. From the combination of these two, many other indicators, such as the various ways of calculating journal impact factor, and h-indices (impact of an author), have been derived. In the description of the theoretical framework, I introduced the sense of constitutive and regulative aspect of PRFS. Typically, it is assumed that bibliometric indicators are not indicators of quality, but rather of performance, impact, usability and other related concepts. In research policy, it is commonplace that bibliometric indicators within PRFS are justified in a general agenda aimed at enhanced quality. Yet, the evidence for such 'enhanced quality' is limited to indicators of performance. This is where in the formulation of research policy the line between quality and performance starts to become somewhat blurry.

An example of such a fusion of concepts is the earlier cited study by Hansen and Lindblad (2010) on publishing within Swedish educational research. This study reports findings also in relation to the PRFS – the Norwegian model – that is used at three universities: Gothenburg, Linköping and Umeå. A detailed description of the Norwegian model can be found on page 33, but the main characteristic of this model is that bibliometric performance is measured employing publication counts. In addition, publications in channels of 'higher prestige' are assigned to a separate 'higher level' – the Level 2. The study by Hansen and Linbdlad (2010) shows that in the three universities on average 27 per cent of all the peer-reviewed journal articles that are included in the model, are assigned to the Level 2.

The share ranges from 25 per cent in Umeå to 30 per cent in Linköping (Hansen & Lindblad, 2010, p.41). Thus it may be said that in doing so a distinction is made between two kinds of bibliometric performance – one of whom (the Level 2 performance) might be associated with higher quality.

A different approach to capture some sense of quality is to employ citations. The meaning of citations has been a topic of wide and intense scholarly debate. An insightful study of the use of citations within the field of educational research was conducted by Michael Hansen (2015). Hansen made an attempt to identify the functions that various citations have. To do so, he analysed a random sample (n=90) of 427 articles in which there were citations of a particular book. His main conclusion was that the function of citations is ambiguous and complex: functions seem to overlap or are hard to identify. In addition, Hansen identified an absence of critical citations and noted that many citations are 'peripheral or ceremonial' (2015, p.14): they are in effect references without analytical contribution.

Sandström and Sandström provide a much more pragmatic view on the meaning of citations. In introducing the Swedish discussion on PRFS that could be used in national research funding system, he refers to 'citations as a proxy for research quality' (Sandström & Sandström, 2009, p.243). In the Swedish system, performance is calculated by employing a citations-based indicator and an indicator based on the amount of external funding (*Prop. 2008/09:50*, 2008). These performance indicators are 'normalised' – the actual numbers are increased by factor of 2 for humanities and social sciences and by factor of 1.5 for natural sciences. Typically such a normalisation is done to acknowledge and somewhat 'compensate' the various publishing and citation practices in different research fields and also the different opportunities to access external research funding. Sandström and Sandström comment that these actual numbers have been 'politically motivated' (2009, p.249) and balanced only the unequal access to funding, but not the various publishing practices. To address this problem, Sandström and Sandström instead propose a 'field-normalised citation rate'. In simple terms this means that first, an average citation rate for field is calculated using WoS data, and then these acquired indexes are used to 'normalise' the actual numerical value that capture the citation rate of a specific field in a specific university.

Such an approach might however not prove appropriate for educational research where the coverage of publications of this field by WoS is rather low (Sivertsen, 2016; Ossenblok et al., 2012). Sandström and Sandström respond to such a potential weakness by arguing that such an approach is accurate when used on national level – the scores are calculated for each university in total: 'We consider field coverage as icebergs; what we see above the water line makes it possible to compute the total activities whether these are in books or institutional reports' (p.247). Evidence for such a claim is a calculation of relative performance indices for a selection of Swedish and Norwegian universities. The calculations compare this field-normalised citations-based index with an index based on the Norwegian model. The calculations indeed suggest that the acquired values are similar. However, commenting on a more detailed comparison of the values calculated for research field, Sandström and Sandström claim that the Norwegian model 'seems to over-estimate the production from humanities and social sciences' (2009, p.248). Moreover, they suggest that the way how the differentiation between 'normal' and 'high quality' is done employing Norwegian model is 'disputable and open to criticism' (Ibid.). They argue that a citation-based indicator is more reliable and better suited for systems where there is a clear 'incentive' for universities that can be expressed in a following way:

> *A university that wants a higher share of the resources should preferably try to strengthen their levels of normalized citations and practise a more selective publication strategy. Salami-slicing of papers will have a negative impact on the citation rate. Authors should try to find their audiences instead of finding the journals with the highest impact. In some areas, such as humanities and soft social sciences, there is a clear incentive for more publications.*
>
> (Sandström & Sandström, 2009)

How educational research performs in such a citation-based research evaluation of educational research in GU can be found in a report of research evaluation of the university that was conducted in

2010 – the so-called *Red10*. This research evaluation was conducted on the basis of a panel review and bibliometric analysis. Bibliometric analysis was conducted using WoS data. Due to the fact that the focus of the study was bibliometric analysis of the whole university, data on publications from educational research were very limited. However, data were reported for the Faculty of Education and also the three departments that were part of the faculty at that time. Educational research was primarily conducted only in one of them – the Department of Education. Here, the RED10 bibliometric analysis for the period of 6 years (2004-2009) reported only 62 publications and 91 citations. It is reported that in terms of 'average field-normalised citation impact', this department performs 'at the world average'. With such normalised citation impact indicators the main problem is that it is hard to interpret them in relation to a more general understanding of research quality. I mean, if RED10 bibliometric analysis indicate that 53.2 per cent of those 91 publications are uncited. This then means 48 publications are uncited, while the other 43 on average receive 2.17 citations. Is it desirable to 'perform at the world average' if it fact means that half of publications of a certain unit are uncited? I would be inclined to say no, since such indicators seem too detached from the actual considerations of what educational research is and what may be regarded as its quality. Yet, such an interpretation are not in line with the view on the quality of research by RED10 Educational research panel.

The panel report of RED10 suggest that the quality of educational research in GU is 'Very good' (for the Department of Education and Special Education), 'Excellent' (for the Department of Education, Communication and Learning) and 'Good' and 'Very good in some aspect' (for the Department of Pedagogical, Curricular and Professional studies (Holmgren & Bertilsson Uleberg, 2011). These assessments are part of a 6-point scale with 'Outstanding' as the higher and 'Poor' as the lowest. The evaluation of educational research can be interpreted in the following way:

> ***Excellent.*** *Research of excellent quality. Normally published so that it has great impact, including internationally. Without doubt, the research has a leading position in its field in Sweden.*
>
> ***Very good.*** *Research of very high quality. The research is of such high quality that it attracts wide national and international attention.*
>
> ***Good.*** *Good research attracting mainly national attention but possessing international potential; high relevance may motivate good research.*
>
> (Holmgren & Bertilsson Uleberg, 2011, pp.31-2)

If these panel assessments are put together with the bibliometric analysis, then it seems that bibliometric patterns discussed above does not correspond to low quality according to peer-assessment, despite my concern of the relationship between 53 per cent uncited publications and the quality of educational research.

As I showed above, evaluative bibliometrics were meant to merely to monitor and describe one or another aspect of research that can be derived from bibliographic data. But this is seemingly no longer the case when bibliometric indicators are employed in already introduced performance-based research funding systems. The origins of PRFS can be traced back to the launch of Research Assessment Exercise in 1986 in the UK (Hicks, 2012). In 2010 there were 12 national PRFS systems, since then the use of PRFS on national and institutional levels continues to expand (Hicks, 2012; Pajić, 2015).

I began this chapter on bibliometrics by referring to the early bibliometric studies in which bibliometrics were assumed to be an unobtrusive method for exploring either scholarly publication or processes in science. Even if in the 1960's it might be said that counting publications or citations might not do much damage, this may no longer be the case with the present widespread use of bibliometric indicators in research funding systems. Various models of PRFS have been constructed with various agendas, but now, after several decades from the origins of evaluative bibliometrics, critical voices keep emerging about the way how the various indicators are used (e.g. Hicks & Wouters, 2015; Rijcke et al., 2015).

A range of studies have explored these intended and unintended consequences by studying publishing patterns (for a recent literature review on the topic see Rijcke et al., 2015). The most straight-forward outcome of PRFS is increase in the number of publications – the outcome that is the intended outcome, if the incentives within research policy are kept in mind.

A study by Moed (2008) explored change over time in publishing patterns in UK in relation to Research Assessment Exercise (RAE) in 1989, 1992, 1996 and 2001. The time-scope of his study spans a period of about 20 years – from 1985 to 2003. The data he used were acquired from Science Citation Index (SCI). Identification with UK articles was bound to authors' addresses – if at least one of the authors indicated her address in UK, an article was assumed to be affiliated to UK.

The choice of data might be somewhat misleading, since in the actual RAE employed data on either all publications of all academic staff (1989, 1992) or data on 4 'best' publications of all 'research active' staff, whereas the study by Moed employed only data on publications indexed in SCI, which includes journals only from natural sciences, medicine and technology. In the same way, among these publications 25 per cent were attributed to non-higher education institutions. Thus, it is possible that the findings from this are not transferable to the context of educational research conducted at public universities. However, these findings are valuable since they cover data from a relatively long period in a country with a pioneering role in the use of PRFS. Findings of the study suggest that over the period of 19 years the mean annual growth rate for UK publication output was +1.5 per cent with standard deviation +/- 2.2 per cent.

Exploration of annual growth rate on year-by-year basis suggested a rather great variation among years. Moed commented that there seemed to be 'timing effects' in the publishing patterns – an increase in the annual growth rate could be observed in the last two years that are covered by each of the RAE. Further, it was suggested that when in 1996 there was a shift in RAE to gather data only on the 'top' publications, there could be observed decrease in the total number of publications, but increase in the share of publications in journals with high impact factors. Moed noted also that it is possible that prior to RAE in 2001, there has been a shift from 'quality' to 'quantity', but this time increasing the number of 'research active' staff. Moed did raise a question on what such changes in patterns would mean, but his conclusion was that more research is needed in order to assess 'whether or not the application of such measures as a research evaluation tool enhances research performance and scientific progress in general' (2008, p.162). In sum, the study seems to suggest that in publishing patterns it is possible to identify changes that correlate the incentives of PRFS, but that some of the changes present unintended consequences.

Similarly, in a longitudinal analysis (1981-1999) of response to the Australian PRFS Butler (2003) suggested that after the introduction of PRFS, there has been a rapid increase in the number of publications. Butler explored responses in relation to quartiles according to five-year average impact factor of the publications within SCI journals. Her findings showed that over a period of about 10 years after the introduction of PRFS, there has been 28 and 15 per cent increase in the top two quartiles, while for the third quartile the increase has been 55 per cent. Exploration of these figures across various research fields indicated that in social sciences there has been even greater increase in the 4rd quartile (more than 60 per cent) and comparatively low for the first quartile (slightly over 10 per cent). The figures for 2nd quartile are also more than 60 per cent, but for the 3rd – a bit over 25 per cent.

Butler commented on this increase with reference to the funding calculation formulas in which the focus is on pure quantity of publications. She argued that, if there is 'no attempt made to differentiate between quality, visibility or impact of the different journals when funding is allocated, there is little incentive to strive for publication in a prestigious journal' (Butler, 2005, p.41). In addition, the study suggested that the response to the national PRFS has varied among universities. Butler noted that this diversity might be dependent on the research distribution schemes that universities employ internally. In her discussion of these findings, an interesting further detail is revealed. After it became apparent that the new PRFS in Australia has led to increase in the *quantity* of publications without any suggestion of other dimensions of these texts, it was proposed to either remove or adjust the

publication component from the funding calculation formula, but it was an organisation representing the universities that argued for the retention of the bibliometric indicator, since it seemed that a great number of universities could benefit precisely from the indicator awarding purely the number of publications (Butler, 2003). Butler's own conclusion was that before changing the actual system, more research is required that would enable a better understanding of validity and fitness-for-purpose of the potentially alternative measures. No argument against the use of PRFS was made, but it is clear that Butler's findings too suggest that unintended consequences arise in bibliometrics-based evaluation.

A study that is more relevant to my topic, more recent and which I have mentioned in earlier discussion, was conducted by Ossenblok et al. (2012). This study explored change in publishing in SSH in Flanders and Norway also in relation to Flemish and Norwegian PRFS. In Norway, PRFS was implemented in 2005 to begin in 2006, in Flanders – the present system that employs regional database is from 2010, before: from 2003, analysis was conducted on basis of WoS data. The study found that there is an increase in WoS coverage for Flemish publishing and the increase in publications in journals channels that are associated with higher quality in Norway. Another note by Ossenblok et al. is that it seems that the Norwegian system has avoided the risk to increase merely the number of publications as identified in Australia (Butler, 2003), while the main conclusion is that 'SSH researchers in Flanders and Norway have responded differently to different PRFSs that have been implemented in their countries' (Ossenblok et al. 2012, p.288). In the discussion of these findings, Ossenblok et al. suggest that the described figures seem to be linked with the incentives identified in PRFS. Implications from these systems are not explicitly discussed, but it is suggested that the actual outcome of these systems and the change in publications might become apparent only in the long term.

A recent study has identified changes in publishing practices at the Faculty of Arts at Uppsala University in Sweden (B. Hammarfelt & de Rijcke, 2015). The aim of the study was to 'depict some of the practical and ethical consequences of output and impact measurement' (Ibid. p.64). The study employed both bibliometric analysis of publications (2006-2013) and also questionnaires for members of staff (n=89; response rate = 26 per cent). The time frame was chosen, as authors suggest, with an intention to enable analyses 'of two "moments of metrics"' – the introduction of the national PRFS in 2009 and institutional in 2011 (Ibid. p.64). The research questions asked in the study are: 'Can we observe changes in publication practices of humanities scholars after the introduction of national model (2009) and the local system for performance-based resource allocation (2011)? And how do scholars themselves reflect on (changes in) publication practices in their discipline?'(Ibid). Drawing on Whitley's (2007) conceptualisation of academic disciplines, the authors of the study explain that the rationale underlying their focus on the humanities is the possibility that due to its specific academic culture the influence of PRFS might be especially visible – either in the form of standardization of publishing practices or in an imitation of publishing that is characteristic of more 'high-status' research fields (Whitley, 2007 paraphrased in Hammarfelt & de Rijcke, 2015). Following such reasoning, the authors expected that the exploration of publishing patterns will show reduction of the diversity in publishing practices.

Findings from the study suggested that the patterns in publishing practices are relatively stable over the period of exploration. However, with respect to publishing language there was decline in Swedish-language publications and increase in the use of English. The questionnaires revealed that the reprioritising of publishing channels is a concern for scholars: 'some scholars see the "devaluation" of the monograph as a threat to the epistemological foundation of their field' (Hammarfelt and de Rijcke, 2015, p.70). For book chapters, a publication type that is rarely discussed, it was argued that the questionnaires seem to suggest that despite the popularity of publishing in anthologies, book chapters have a rather low status. Another finding from the study is an increase in peer-reviewed publications from 20 per cent in 2006 to 49 per cent in 2013. Noting the limitations of the study, Hammarfelt and de Rijcke (2015) pointed out that in case of various overlapping PRFS it is hard to disentangle the influences of one particular system. Yet, the focus on the 'moments of metrics' – the time period when a PRFS is either introduced or changed – may be a way to gain some sense of the influence of a

particular system. This suggestion corresponds with findings from the UK by Moed (2008) and from Australia by Butler (2003).

The main conclusion from the study on publishing practices in Uppsala University is that it is not possible to 'make the causal claim that the implementation of evaluation models at the national and local level is solely or even mainly responsible for these changes. It is however hard to ignore that these developments are perfectly in line with the incentives in both the local allocation model at Uppsala University and in the national model: To publish for an "international and scholarly audience"' (p.74). As final thought it is suggested that it may be that effects of bibliometrics-based research evaluation on the actual funding that is allocated within these systems are of no great significance, as opposed to the clash between incentives that are induced by PRFS and the particular characteristics of the various research fields within academia.

Changes have been identified also in publishing patterns of educational research within GU (Foss Lindblad & Lindblad, 2016a). The authors report that from 2004 to 2014 the total number of publications has increased from 346 to 670 publications a year, while the number for peer-reviewed journal articles has risen from 41 to 199 publications a year. Does this suggest an increase in quality? Foss Lindblad and Lindblad suggest that the increase in the number of articles might be problematic. In their argument a link is drawn to the PRFS that is used at the Faculty of Education: 'The simple assessment of research quality in terms of the assumed prestige of publication channels (the Norwegian model) does not show a larger share in high-quality journals' and, in addition, 'research communication analyses present a scattered picture with fragmented communication and a lack of core journals according to current communication patterns' (p.221). It is proposed that such 'core journals' are 'vital for the intellectual and social recognition of a field of study' (Ibid. p.222).

A Spanish study offers a somewhat different angle on the discussion of the unintended consequences of the use of bibliometrics-based research evaluation. Its authors suggest that in fact, there are no means to link publishing patterns with PRFS (Osuna, Cruz-Castro, & Sanz-Menéndez, 2011). After having analysed the validity of several other studies that have claimed causal relationship between PRFS and growth in publications, Osuna et al. (2011) claim that any growth may sooner be attributed to increased expenditure in research and development or to increased number of researchers and maturation of research community, but *not* to the introduction of PRFS.

It may be that Osuna et al. (2011) and Hammarfelt and de Rijcke (2015) are right: causality between PRFS and change in publishing patterns might not be a phenomenon that can be empirically captured. However, recalling the discussion in the section on social epistemology, it is nevertheless possible to *theorise* changes in publishing patterns. Going beyond the search of causality, and turning instead to a social epistemology that establishes a metaphysical frame for understanding bibliometrics-based research evaluation, makes it possible to simple state, firstly, that there is change in publishing patterns. Then, this change may be taken as a starting point for further exploration of the range of possible meanings that may be attached to it, focusing on the field of educational research. In order to do that I return again to a literature review of educational research, this time focusing on the various ways the notion of research quality is understood.

Elusive Quality of Educational Research

A characteristic of educational research is not only the wide diversity in the ways the quality of educational research can be understood, but also on what educational research is. At the beginning of my studies here at the University of Gothenburg, I was introduced to a book by Anderson & Arsenault (2005) 'The Fundamentals of Educational Research'. In this book, it is said that educational research is 'a disciplined attempt to address questions or solve problems through the collection and analysis of primary data for the purpose of description, explanation, generalization and prediction' (p.6). Such a definition of educational research characterises the so-called positivism approach to educational research. These same authors argue that 'research reflects the values, beliefs and perspectives of the researcher' (Anderson & Arsenault, 2005, p.2). I would most certainly agree with the latter: the diversity that can be seen in the ways educational research is understood and done can well be

explained with reference to the plurality in values among researchers. The particular, positivist definition of educational research introduced by the authors themselves is precisely such an example of how an understanding of educational research mirrors the values, beliefs and perspectives of the researcher. In other words, sets of values, beliefs and perspectives may be seen as characteristics of particular research paradigm.

Widespread paradigms that can be identified in educational research are, as already-mentioned, the positivism/post-positivism paradigm, the interpretivist/constructivist paradigm, and the emancipatory and deconstructivist paradigm (Lather, 1992). Lather suggests that each of these paradigms distinctively vary in *the aims* of research, which may be to predict, to understand, to emancipate, or to deconstruct. Consequently, all the further considerations how research is conducted and how its conduct evaluated are derived from these general aims. For the purposes of this study, further details on these paradigms might not be necessary. It is sufficient to acknowledge that such diverse aims are possible in educational research, and that they are central to the various ways of understanding the quality of educational research.

Yet a brief further comment on this diversity might be apposite, taken from a text by philosopher of education David Bridges (2009). He reflects upon the assessment of the quality of educational research and points out that the quality of educational research is not a quality that can be detached from the context of either the text or from the reader. He reminds that in order to recognise the value of a particular text one has to understand how various kinds of research are done and what kinds of knowledge claims already exist in the context. Bridges suggests that only in this way it is possible to identify the quality of educational research. He continues with the suggestion that proxies of quality, such as the number of publications in a particular journal or the number of citations, are merely correlates that may or may not indicate the quality of a particular text or, on aggregate level, the quality of research in a specific university or research field. He concludes: 'Quality in research is not reducible to a single set of values, nor representable by a single set of measures on a scale. In making qualitative judgements we have to find a way to hold a plurality of values in our minds at once and to discover such as are appropriate in the object under scrutiny' (Bridges, 2009, p.513). This does seem to be a more flexible way to view the quality of educational research. Yet there is a danger of simply avoiding the challenging meta-discussion on what educational research is and what might be the ways to conceptualise quality in a context of such diversity.

There are more obstacles for such a meta-conceptualisation. In addition, great differences in the understanding of educational research can be identified in various national contexts. Originally, diversity of this kind much characterised the so-called Anglo-American sense of educational research. In this tradition, educational research is seen as a multi-disciplinary research field that is rooted in three academic disciplines: sociology, psychology and history (Depaepe, 2002). Yet, today the trajectory of educational inquiry spans a much greater number of disciplines – also including economics, geography, philosophy and many others (Furlong & Lawn, 2011). This multi-disciplinary way of understanding educational research contrasts with a continental tradition, where educational research is homed in a single academic discipline called 'Pedagogy'.

In a historical sense, the primary goal of educational research within Pedagogy discipline has been to provide a single scientific base for teacher education. Its focus has consequently been more on understanding the various phenomena surrounding schooling and education. I should note that such sharp distinction between an Anglo-American, multi-disciplinary educational science research tradition and a continental research tradition of pedagogy as single discipline might not hold in the present contexts. Traditions have transformed, blended and acquired new forms; and yet some imprints of the original orientation might be of importance when exploring the case of Swedish educational research.

The contribution of educational research to knowledge is complemented by a strong sense of its social value. This social aspect of the quality of educational research may be grasped by thinking of the social actors who are at 'the receiving end' of knowledge acquired by educational research. Be it

teachers, policy-makers or simply curious members of society, each of them has a more or less distinctive understanding of the aims of educational research and what knowledge is expected from it.

An example of this diversity of understanding quality in action is the debate about the certainty of knowledge claims. It is argued, that for teachers and education policy-makers the main expectation from research is that it is practically applicable and, most importantly, *certain* knowledge. Among teachers and policy-makers it is common to expect that research will provide answers on what decisions should be made or what outcomes can be expected given these decisions. In contrast, within the research community it is generally regarded as self-evident (except in the notable case of evidence-based educational research) that there are no means by which such certainty might be established (e.g. on relevance to practice - Winch et al., 2015; to policy - Welch, 2015). Consequently, whenever such an expectation of certainty guides a judgement about the relevance of educational research that offers no such certainty, then that educational research will likely be deemed irrelevant and hence *lacking quality* (see analysis of such a critique in Oancea, 2005; also Pring, 2015, Chapter 1 and 9).

There have been various attempts to address this mismatch between expectations and actual possibilities, either by calling for, for instance, more teacher training (Winch et al., 2015) or for more rigorous research that would meet the criterion of certainty (e.g. Torgerson & Torgerson, 2001); but none of these ideas seem to silence the continuous debate on the quality of educational research (see Oancea, 2016; Reid et al., 2014).

Philosopher of education Gert Biesta (2014) has recently theorised the various ways of engaging in the discussion on the quality of educational research and the political implications that may follow. Biesta constructs his arguments by drawing from theoretical insights by philosopher of science Gerard de Vries (1995)[1]. De Vries proposes a theoretical model aimed at illuminating the role of research for practice that distinguishes between a technological and a cultural role. If the technological role is understood in a know-how sense, then the cultural role emphasises pluralism in interpretations. Further, de Vries suggests that this distinction means that these roles are mutually exclusive. The same knowledge claim can be one or another, or both in more or less equal proportion, but never be one and the same thing. In addition, de Vries too highlights the situatedness of knowledge practices. Biesta describes this as follows: 'In those cases in which there is a strong consensus about the aims of education or, to put it differently, where the aims of education cannot be questioned, the only 'possible' role for research seems to be a technical role. When, on the other hand, such a consensus does not exist, there is a possibility for research to play a cultural role by providing different interpretations of the situation' (Biesta, 2014, p.397). Here the role of democracy enters discussion of research quality as given by relevance. Biesta, referring to de Vries, suggests that a 'democratic society is . . . characterized by the existence of an open and informed discussion about problem definitions and the aims and ends of our educational endeavours'. Consequently, if the discussion about the quality of educational research is dominated by, for example, the evidence-based sense thus hindering the potential cultural role, it can be taken as 'a worrying sign from the point of view of democracy' (Biesta, 2014, p.398).

Similarly, philosopher of science Bueter (2015) argues that 'epistemic trustworthiness of science is promoted more by pluralism of non-cognitive values rather than by their exclusion'(p.18). Exploring the plausibility of the long-pertained ideal of value-freedom within knowledge-making, Bueter suggests that 'it is not a feasible goal for science to uncover the whole truth about the world. Instead, science aims to provide significant truths', thus rendering this knowledge-making enterprise democratic (Ibid., p.22). Hence the plurality of educational research and the consequent plurality in conceptualising the quality of research of this kind might turn out to be beneficial qualities of research in their own right, a point reflected also by Bridges (2009).

[1] Biesta (2014) refers to Vries (1990) , yet all the other bibliographic information suggests that he means Vries (1995).

Swedish Educational Research

In Sweden, educational research has been linked with pedagogy (*pedagogik*) as a single research discipline as in the continental tradition, but in the recent decade top-down reorganisation has effected an inter-disciplinary research field of education sciences (*utbildningsvetenskaperna*), thus introducing a collision between the various ways of understanding the relevance of educational research (Foss Lindblad & Lindblad, 2016b).

In a historical sense, the role of educational research in Sweden was tied closely with teacher education, whereby the task of pedagogy was to provide 'philosophical rationales for goal-setting and the psychology of student learning' (Foss Lindblad & Lindblad, 2016b, p.70). Similarly, Englund argues that at the first half of the 20th century in Sweden 'the object of study of educational science was the basic conditions at the individual level for development and learning' (Englund, 2006, p.383). Such a close relationship between educational research and education policy has characterised most of the history of 20th century educational research in Sweden. It has been suggested that due to this reason Swedish educational research acquired characteristics of instrumentalism and consequently produced some 'blind spots' in the choices and conceptualisations of research problems (Englund, 2006; Vislie, Popkewitz, & Bjerg, 1997). As a result, the question of the relevance of educational research is in Sweden too closely linked to the academic culture and its particular relations with the various social institutions involved.

In the 1970's, educational research in Sweden was complemented by more critical studies that were inspired by French and British sociology (Englund, 2006). However, it is not clear how critical these studies have been, since there is a contradicting suggestion that these so-called critically oriented studies were incorporated in dominant arguments in social policy, thus subscribing to the same instrumental sense of educational research (Vislie et al., 1997). These insights by Vislie et al. (1997) come from a report of an international evaluation of educational research in Sweden carried out in the 1990s. A crucial insight from this report was the sense of diversity that characterised the Swedish educational research at that time (Rosengren & Öhngren, 1997, p.17). In terms of conceptual frameworks and methods of data collection and analysis common in Swedish educational research, the report names action research, ethnography, psychometric methods as well as experimental methods. Research focus covers child development, learning and classroom interaction, educational systems and policy and many other themes. In terms of education level, research is conducted on all levels, from pre-school to higher education as well as special and vocational education, and teacher training.

Soon after the 1997 report on the Swedish educational research, the public research funding system for educational research was changed from a direct state funding to a competition-based funding (Foss Lindblad & Lindblad, 2016b). In this new configuration, competition for funding educational research was introduced among the various disciplines that now were legitimate potential grantees of educational research funds. Foss Lindblad and Lindblad (2016) comment that one of the reasons for this change was that the research conducted within pedagogy was regarded as 'being too politically informed and lacking scientific quality' (p.73). Change in the way educational research is formally defined, was complemented also by a general liberalisation of Swedish HE and the introduction of PRFS in 2009. It has been suggested that 'the relevance of educational research' 'was lost' and with these changes and academics 'were forced to navigate in a multi-disciplinary landscape…all competing over scarce resources and measured by different performance indicators' (Foss Lindblad & Lindblad, 2016, pp.74-5). Such commentary may indicate the tension between the ways how educational research is conceptualised within Pedagogy discipline and how it may be seen when educational research is understood as multi-disciplinary research field. It may be the case that the seeming lost relevance is rather an indication of a shift to a pluralistic understanding of educational research. Although, from the perspective of Pedagogy discipline it may seem as a loss of the historical link between educational research and the national education policy-making that has rendered this kind of educational research relevant.

A somewhat contradictory argument on the relationship between educational research and education policy is provided by Englund (2006) in relation to the education reforms carried out in Sweden in 1990s. His suggestion is that, within that public debate, educational research was simply ignored by education policy-makers. Instead, it was the public media that turned out to have a voice in decision-making. A parallel discussion on educational research and its quality is taking place in relation to changes made in teacher training (Beach, 2011).

Without going into further details on the content of each of these moves in Swedish educational research, it does seem to be the case that the recent two decades are a sort of emergence of a new era, in which there is no clear understanding, neither among researchers nor among policy-makers, what educational research is and what its goals are. In this context, it is quite impossible to speak of a uniform sense of research quality. Not only within the various disciplines now joined-up as educational sciences, but also within the borders of pedagogy perceived as one discipline, the sense of quality is likely to find various radically different conceptualisations. If this is indeed the case, then the question becomes, is there any particular loss in such plurality of conceptualisations of research quality? Is it truly undesirable for teachers, policy-makers or society at large that within educational research one can find both quantitative, post-positivism minded scholars working in the same hallways as, say, feminist and Marxist-inspired scholars, all of whom have a varying sense of the research qualities they pursue?

My own inclination is to respond that no, such plurality is precisely a strength of educational research. In settling for this position I draw from the earlier introduced arguments from Bueter (2015), Biesta (2014) and Bridges (2009). How such reasoning manifests in a context of PRFS, to find this out is the task of the next section where I render the just introduced consideration through the lens of social epistemology.

Theorising Publishing Patterns

In this section I develop further my approach to conceptualising change in publishing patterns that occur in a context of PRFS in relation to the quality of educational research. Before I proceed, I return to a few points from the section on social epistemology. There, I introduced a particular way to theorise the potential benefits and losses whenever change occurs. I referred to regulative and constitutive aspects of PRFS. In simple terms, this distinction refers to a subtle conceptual change whereby notions of research quality are replaced with single a notion of bibliometric performance. If there is change in publishing patterns, that change can be conceptualised as change in the codified practices within the research community. This is the point that needs to be extended further.

If change in publications is seen as change in codified practices within a research community, there is no need to suggest that PRFS makes me change my beliefs about the quality of educational research. It is sufficient that I act as if I had changed my beliefs merely by acting in accordance with PRFS, thus contributing to what would appear to be an ontological change (a change in my foundational conception of what constitutes good research) on the collective level—since under common subscription to PRFS, I and others are encouraged to act differently in the very same direction.

Fuller suggests that whenever occurs such a collective change in codified practices, there may emerge knowledge effects. Knowledge effects refer to the interactions between the knowledge codified in a form of, say, books and, for example, teachers. Books that teachers might regard as valuable might or might not coincide with the worth that is attributed to such books in PRFS. But if researchers begin to adjust their publishing according to PRFS, then such behavioural change can lead to a situation where a group of people who had relied on a particular source of knowledge are deprived of it. This is how it is possible to theorise changes in publishing patterns in relation to the quality of educational research that occur in a context of PRFS.

Another example of thinking about knowledge effects may be the present configuration of academia being related to the original idea of a modern university – the Humboldtian ideal. The traditional Humboldt university was established on a principle of unity between research and teaching with

precisely such an innovation rationale: knowledge that is just acquired is immediately distributed to students. Fuller (2003) has called this unity between research and teaching the creative destruction of social capital within universities: 'as researchers, academics create social capital because intellectual innovation necessarily begins life as an elite product available only to those on "the cutting edge". However, as teachers, academics destroy social capital by making the innovation publicly available, thereby diminishing whatever advantage was originally afforded to those on the cutting edge' (p.107).

In a typical university today, research is alienated from teaching. In the earlier cited RED10 report it is suggested that at the Faculty of Education there is differentiation between research time and time for other responsibilities (such as teaching). It is also indicated that the amount of allocated research time varies among the various professional categories. In the Department of Education and Special Education, for example, professors have about 50 per cent (range 16-97 per cent) of time allocated for research, but associate professors about 40 per cent (range 6-97 per cent) (Holmgren & Bertilsson Uleberg, 2011, p. 10). I do not know what the present situation at the Faculty of Education is, but if it remains similar to what was reported in RED10, then it may be the case that teaching, which under the Humboldtian ideal would prevent knowledge from acquiring power and contribute to the spread of knowledge, acquires a status of punishment for failing to acquire research funding. Considerations such as these seem to remain entirely unexplored within educational research, but let me first bring together the thoughts derived from arguments found in the literature.

Gaps in the Literature

In the literature, two major gaps can be identified – a theoretical gap and an empirical gap. In empirical explorations of changes in knowledge practices on the basis of publishing, changes are commonly expressed as numbers of publications in data that are primarily descriptive and that are either under-theorised or not theorised at all in relation to the quality of educational research. For example, the evidence suggests that there is an increase in the number of the most highly prioritised type of publication within PRFS, the number of peer-reviewed journal articles published in English (Foss Lindblad & Lindblad, 2015; Hansen & Lindblad, 2010; Kyvik, 2003). Does this mean that an increase in the number of publications is an increase in the quality of educational research? And how do such PRFS data relate to the fact that an increase also occurred before the introduction of performance-based research funding systems (2009)? In the absence of detailed theory-driven analysis of these patterns, answers to such questions inevitably remain stuck at the level of suggestiveness. Yet publishing patterns within educational research are likely to continue to change, and understanding changes that may follow from the use of PRFS in particular seems crucial for grasping their implications for the quality of educational research.

Earlier, I have referred to a few theoretically oriented studies that in one way or another theorise problematic aspects of the quality of educational research. There, it is common to assume that the answer to many of the problems of educational research can be found in more education (Winch et al., 2015) or better organisation of research (Foss Lindblad & Lindblad, 2016b), but none of them seem to be putting in question whether PRFS could itself be a source of concern with respect to research quality. At the same time, studies on educational research take the status of research as a necessary and trustworthy process leading to new knowledge for granted and so have little to say about research as politics where a distinction is made between more and less powerful assumptions of quality and relevance.

By contrast, in studies that do acknowledge the need to be aware of the use of PRFS and do take a more reflective stance towards research, discussions do not consider specifics of the field of educational research. Hence, the direction that I wish to take in this study is to explore the implications of PRFS for the quality of educational research, but without falling into a prior assumption that the worth of research can be taken for granted. The more critically reflexive direction that I am thereby taking should encourage discussion about the goals of educational research and question the taken for granted assumption that more research and more publications would reflect a rise in the quality of that research.

I am now ready to set out my particular case: educational research at the Faculty of Education (FEdu) of the University of Gothenburg (henceforth GU) in Sweden, over the period 2005 to 2014. Characteristic of this period is the introduction of PRFS on the national and university level midway, in 2009 (*Prop. 2008/09:50*, 2008; University of Gothenburg & Universitetslednings kansli, 2008). Thus, like in the study by Hammarfelt and de Rijcke (2015), I will capture one of the 'moments of metrics', but here specifically of educational research.

I propose to study my case through a theoretical and an empirical part. In the empirical part, the educational research is explored on the basis of its publishing patterns. The focus is on change in the number of peer-reviewed journal articles in English, in comparison to other publication types. My choice in favour to this particular type is purely pragmatic: peer-reviewed journal articles are the type of publications that is most discussed in the literature on publishing patterns, while within the field of educational research peer-reviewed journal articles have not traditionally been the dominant publishing channel.

As my literature review shows, existing studies have already noted that such rise in this publication type exists (Foss Lindblad & Lindblad, 2016a) and this rise has already been identified before the introduction of PRFS in GU and Sweden (Hansen & Lindblad, 2010; Holmgren & Bertilsson Uleberg, 2011). However, the level of detail in the empirical data within these studies is overly general. Either the data are provided only using publication counts according to publication types only (Foss Lindblad & Lindblad, 2016a), the inclusion criteria are questionable (Hansen & Lindblad, 2010), or only a small portion of educational research publications are studied (Holmgren & Bertilsson Uleberg, 2011). Consequently, it seems problematic to interpret such figures in relation to the introduction of PRFS at GU and the quality of educational research. To address these gaps, the empirical part of this study explores changes in publishing patterns in relation to the PRFS used at FEdu, the categorization of publications used in GU and the use of language. I expect that this will provide a much more fine-grained basis for theorising changes in the publishing patterns – the second part of this study.

The Aim and Research Questions

This study aims to theorise implications to the quality of educational research that may be identified in the changes of publishing patterns in a context of PRFS. This aim is achieved by addressing the following two research questions – one empirical (1) and one theoretical (2):

1. What change (if any) can be identified in publishing patterns in educational research (2005-2014) before and after the introduction of the GU performance-based research funding system in 2009?

1.1. What change (if any) can be identified in the number of peer-reviewed journal articles in relation to:

   a. The categories used at the funding system and language?
   b. The categories used at the funding system and other types of publications?

1.2. Is there compliance in publishing patterns in GU educational research (2005-2014) with assumptions implicit in the GU performance-based research funding system introduced in 2009?

2. Given the available literature on the topic and the actual publishing patterns, what likely implications (if any) does the introduction of the GU performance-based research funding system have on the quality of educational research?

Method

The empirical research object of this study is publishing patterns of educational research. I explore change in publishing patterns employing bibliometrics – a quantitative approach to study publication data. The design of this study is guided by considerations rooted in the theoretical framework of this study – social epistemology. I first refer to considerations regarding the operationalisation of concepts. Then, I proceed to a more detailed account of the description of data, approach to data analysis and their interpretation, and also ethical and legal considerations.

Operationalisation of Concepts

One of the main principles of Fuller's (2002) social epistemology that applies to research design is that the explored categories are derived from the context of the study. In this study, that context is the University of Gothenburg. I do not introduce definitions, but I search for categories that relate to the concepts I am interested in.

The concept *'educational research'* is defined as research conducted in those administrative units within the Faculty of Education at the University of Gothenburg where (1) research is primarily on education and (2) those who have a reasonable share (more than 5 units a year) in the total amount of publications of the Faculty registered within a database 'Gothenburg University Publications' (GUP).

Applying such a definition, the units included in the study are (1) the Department of Education (for the period from 2005-2010), (2) Department of Education and Special Education (2010-2014), (3) Department of Pedagogical, Curricular and Professional studies (2010-2014), (4) Department of Education, Communication and Learning (2010-2014) and (5) The Linnaeus Centre for Research on Learning, Interaction and Mediated Communication in Contemporary Society (2006-2014).

A unit that is not included in the study according to the first criterion (above) is the Department of Food and Nutrition, and Sport science. There, along with educational research, research is conducted that is characteristic to other research areas, for instance, medicine. Applying the second criterion, the excluded units are 'Pedagogical Development and Interactive learning' (PIL), and the Centre for Educational Science and Teacher Research (CUL). This set of administrative units serves as a criterion for inclusion of bibliometric data.

The concept *'publications'* is defined as the types of publications that are recognised by the GU database GUP. The complete classification of publications within GUP amounts to 25 different types, but in the data used in this study only 21 types were identified (for definitions of these types see Appendix 1). Further details on the outcome variables that count the number of publications are described in the next section.

The '*performance-based research funding system' (PRFS)* on which this study is focused is the Norwegian model that is used to calculate share of the funding that from GU is allocated to the Faculty of Education. Since this model is central to this study, I provide detailed description of this model in a separate chapter of this thesis (begins on page 33).

Change is a concept that in this study I express with a numerical value. I identify change in three ways: change in *quantity*, change in *speed* and change *towards PRFS* or – in other words - *compliance.* The calculations that I use for quantity and speed I will describe in the section on analysis, but with *compliance*, I mean change in only those publications as might be expected precisely given the priorities set in PRFS. It means that I acknowledge the fact that I cannot access beliefs of authors of publications whose data are analysed in this study, but I can view these data from a perspective of a research policy-maker who expects that the use of PRFS will lead to increased quality of research and the evidence for such an increase can be found in bibliometric data. Hence if there is increase in those types of publications in which I *expect* increase, then I can interpret it as *compliance* with the assumptions of the particular PRFS despite the fact that there can be a range of

objections that can be raised against such reasoning. These assumptions in which compliance might be expected I will introduce after the description of the $PRFS_{GU}$ on page 38.

Data and Descriptive Statistics

The data used in this study are bibliometric data retrieved from a database Gothenburg University Publications (GUP). Bibliometric data were provided by the GU bibliometrician Bo Jarneving on 6 and 13 October, 2015. The data contained a list of all publications affiliated with the Faculty of Education between years 2005 and 2014. The format of data: two Microsoft Excel spread-sheets (A and B). One contained all publications; the other contained a subset of publications that were included in the calculations for research funding allocation purposes. The unit of analysis is a publication.

Data contain the following characteristics of each publication: (1) the identification number of a publication (A;B); (2) the year of a publication (A;B) ; (3) the title of a publication (A; B); (4) the name of the author of a publication (A only); (5) the institution affiliated with the publication: Research Environment (2005-2010), Department (2010-2014) (A ; B); (6) language of the publications (A;B); (7) type of the publication (A;B); (8) level assigned to the publication in the Norwegian model (B only). For further analysis of the data, the two spreadsheets were merged using the identification number of the publications.

The data on the level of a publication have been retrieved from the list of publishing channels and their assigned levels that has been compiled in January, 2015 by Norwegian Centre for Research Data. Details on this process of assigning levels will follow.

For further analysis, a separate dataset was produced with data about publication characteristics. The dataset contains 6189 cases. This number includes duplicate entries: publications that are affiliated with more than one author or more than one department. The number of unique publications without duplicates is 4134. This set of publications is further reduced to 4090 publications. Description of the reasoning that underlies this reduction follows shortly.

Each publication is described with 5 variables: their identification number, publication year, publication type, publication language, level in the bibliometric model (see Appendix 3Appendix 3). The original set contains 21 categories of publication types but for further analysis it is reduced to 16 categories. This reduction is conducted by setting up exclusion criteria and combining some of the categories (Appendix 2). Types of publications that are not part of the data of this study are ‘Artistic research and development project’ (n=1; 0.02%) and ‘Other’ (n=58; 0.94%). Three pairs of categories are combined into three variables: this applies to books, edited books and book chapters that in original form are differentiated between peer-reviewed and non-peer-reviewed. It was confirmed by GU bibliometrician Bo Jarneving that this differentiation has been introduced only after 2012 (B. Jarneving, personal communication, April 20, 2016). To avoid misinterpretation, the differentiation between peer-reviewed and non-peer-reviewed is removed. This final dataset is the earlier mentioned reduced dataset that contains data about 4090 publications. Descriptive statistics of the dataset can be found in the Appendix 5 and Appendix 6.

The original data contained publications in 11 different languages. However, further analysis was conducted with three categories for language: English, Swedish and other. In the original dataset there were categories for other languages, but due to the fact that these publications accounted for only 3 per cent of the total number of publications, these categories were merged in the category ‘other’. The original categories and the number of the respective publications were the following: Norwegian (n=35), German (n=21), Danish (n=17), French (n=14), Spanish (n=4), Italian (n=1), and other (n=20). In the final dataset, more than half (58 per cent; n=2383) of the publications are in English, whereas 39 per cent (n = 1595) of publications were in the national language, Swedish. The publications written in other categories account as noted for 3 per cent (n=112). Due to the focus on the language only for peer-reviewed articles, these figures are the only ones for language that I provide in this paper in relation to all the publications.

The variable 'Level' accounts for the level assigned to a publication in the PRFS of the Faculty as described in an earlier section of this thesis. From the dataset of 4090 publications, 2505 or about two thirds of the publications are included in the analysis conducted for the research funding allocation purposes. For this reason the data on the assigned level of quality are available only for this subset of 2505 publications.

Earlier, I mentioned that data on publications will be counted as full-counts, without fractioning according to the number of authors or institutions. The reason for this choice is the lack of data on the number of authors for the whole dataset used in this study. By doing so it is no longer possible to acknowledge the extent to which the increase in the number of publications can be linked with increased collaboration with other institutions.

The dataset was produced in Microsoft Excel. The final dataset, once prepared for actual analysis, was exported to the Statistical Package for the Social Sciences (SPSS), acquiring the dataset in .sav format.

Analysis and Interpretation of Data

Figure 1 (page 31) presents the conceptual scheme of this research design, which contains empirical and conceptual level. The empirical level is the level of publications. The design of this study is longitudinal: the data analysed in this study cover a period of ten years from 2005 to 2014 (included). As I have mentioned earlier, in 2009 PRFS was introduced in GU. Therefore, the time-frame of the study enables the exploration of change in publishing patterns *before* (2005-2009) and *after* (2010-2014) the introduction of $PRFS_{GU}$. Hence I can discuss change in publication patterns in relation to PRFS.

The analysis of data is conducted employing descriptive statistics. I use cross-tabulation to explore the relationships between the number of publications in relation to various categories (type, language, level in PRFS). In addition, I calculate percentage rate (PR) and growth rate (C) for the number of publications (see Equation 1 and Equation 2) as it is done in studies by Ossenblok et al. (2012) and Jarneving, Henning and Sandberg (2012).

Equation 1 Percentage rate

$$PR = \frac{(P_b - P_a)}{P_a}$$

PR – percentage rate rate

P_a - number of publications for the earlier period

P_b - number of publications for the later period

The percentage rate (Equation 1) I calculate in order to identify the difference in the total number of publications of a particular type for the whole period (a) before (2005-2009) and (b) after (2010-2014) the introduction of PRFS. Equation 2 I employ to calculate annual average growth rate for the number of publications for a period (a) before (2005-2009) and (b) after (2010-2014) the introduction of PRFS. For those categories in which the numerical value for the first year of the period (2005 or 2010) is '0' (zero), the growth rate is not calculated.

Equation 2 Annual average growth rate

$$C(\%) = (\sqrt[n]{\frac{N_{y2}}{N_{y1}}} - 1) \times 100$$

N_{y1} number of publications for the year 1

N_{y2}	number of publications for the year 2
C	growth rate for year 1 in relation to year 2
n	the number of years

Statistical analysis is conducted with Microsoft Excel and Statistical Package for Social Sciences (SPSS). The graphs are created in Microsoft Excel.

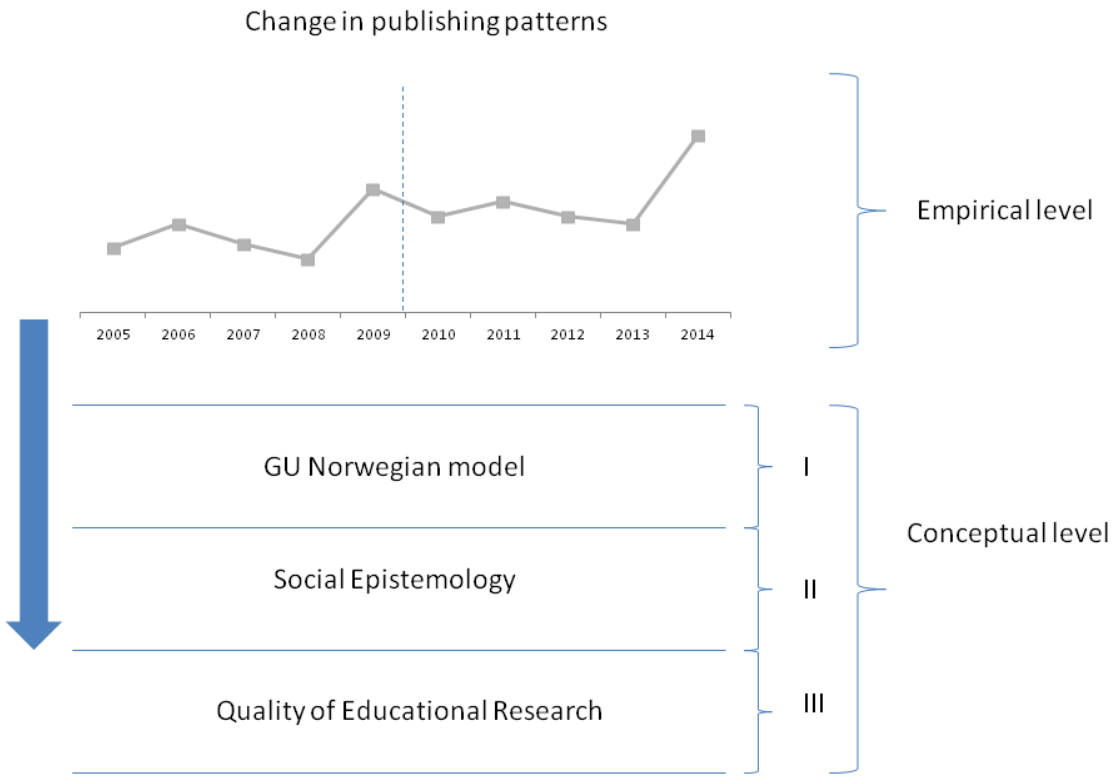

Figure 1 Conceptual scheme of research design

In order to theorise implications to the quality of educational research that may be identified in these changes of publishing patterns that occur in a context of PRFS, I introduce three additional conceptual levels (Figure 1). The first one is the level of $PRFS_{GU}$, the Norwegian model. It means that within this level of interpretation, the change in publishing patterns is analysed through the lens of GU Norwegian model. Details of this model will follow in a separate chapter (begins on page 33).

The design of the Norwegian model and its use in GU is the next context that I analyse in order to identity potential implications to the quality of educational research. To do so, I identify assumptions about *goals of research* implicit in the design and the use of $PRFS_{GU}$. I search for these assumptions in descriptions of the PRFS and the policy documents describing the purpose of the implementation and the use of the system. The sources are scholarly articles in which the Norwegian model has been described and the so-called *grey literature:* reports, national statistics, meeting minutes, presentation slides as well as GU website and personal communication with GU bibliometrician Bo Jarneving.

Next, these assumptions about goals are related to educational research. To do so, I follow the reasoning of social epistemology. I attempt to *falsify* the goal(s) that I have identified as present in the context of PRFS by relating them to insights on the quality of educational research, presented earlier in the chapter entitled 'Bibliometrics and the Quality of Educational Research' (page 10). In a way, I

explore whether claims about research goals as is implicit in the Norwegian model holds also in a context of educational research. This is where the normative part of social epistemology manifests. The interpretation of findings through both the Norwegian model and the social epistemology I present in the chapter devoted for discussion (page 47).

Ethical and Legal Considerations

In this study, bibliometric data are analysed. Therefore ethical issues are related to the use of these data. I have derived insights on addressing ethical issues from Miles and Huberman (1994, 290-7 in Punch, 2005, pp.277-8) and from guidelines in 'Good research practice', a text issued by the Swedish Research Council (Vetenskapsrådet, 2011).

A general principle that guides the ethical consideration within this study is that 'the research criterion' that aims at advancement of knowledge as an interest is balanced with 'the criterion of protection of the individual' (Vetenskapsrådet, 2011, p.18).

The legal framework regarding the ethical issues that I have consulted is 'The Act concerning the Ethical Review of Research Involving Humans' (*'Lag (2003:460) om etikprövning av forskning som avser människor',* ERH henceforth), The Ministry of Education and Cultural Affairs, 2003. It is an English summary of ERH.

Given sections 3 and 4 within ERH that define the characteristics of research that would fall under the legal power of ERH, I have assumed that the study as I conduct it does not require an ethical review in the sense of ERH. Therefore, the ethical considerations I refer to now are guided instead by general research ethics (Punch, 2005).

In this study I conduct a secondary analysis – I analyse already collected bibliometric data. Nevertheless, these data are about people, authors of publications within educational research. The data I am exploring are related only to the public actions of academics: their being an academic in a public institution and publishing as affiliates of that public institution. Thus, the data are regarded as public (Offentlighets- och sekretesslag 2009:400; Sverige & Regeringskansliet, 2009; Personuppgiftslag 1998:204; Sverige & Justitiedepartementet, 2006) and fall under the legal power of the'principle of public access to information' (Sverige & Regeringskansliet, 2009, p.7). According to this principle, all members of society 'are entitled to receive information about state and municipal activities' (Sverige & Regeringskansliet, 2009, p.7), thus data about university publications are such information that is available to public access (Ibid., pp.7-11).

Despite the fact that these data are available for public access, the bibliometric data acquired will be used only for the purposes of this study. The data will not be shared and data that can lead to identification of individual people will not be reported.

On completion of the study, the results obtained through the use of the data will be reported to the Dean of the Faculty of Education so that further distribution of the findings to the authors of the publications included in the data may be arranged.

Now, I proceed to a description of the context of the study and the Norwegian model in its original form as well as in the way it is used in GU.

Educational Research and Bibliometrics in GU

In HE in Sweden, there are 14 public-sector universities, 17 public-sector university colleges (Swedish Higher Education Authority, 2014). The University of Gothenburg is the third largest university in Sweden in terms of full time equivalents of teaching and research staff (n=2563) (Ibid.).

Currently, most of educational research in GU is conducted at the Faculty of Education; however, some research is conducted also in other Faculties – for example, at the Department of Sociology at the Faculty of Social sciences. The Faculty of Education at the University of Gothenburg has a rather recent history. It was founded in 1990. During the first part of the period of interest within this study, from 2005 to 2010, the Faculty consisted of 3 main Departments: Department of Education (*Institutionen för Pedagogik och Didaktik)*, Department of Work Science (*Institutionen för Arbetsvetenskap*) and the Department of Food, Health and Environment (*Institutionen för mat, hälsa och miljö)*. In addition, in 2006 The Linnaeus Centre for Research on Learning, Interaction and Mediated Communication in Contemporary Society (*Linnécentret for forskning om lärande*) was established. During the period there were also some smaller administrative units, which are not accounted for here due to the fact that their contribution to the scholarly communication of the Faculty either is not observed, or is minor (less than 5 publications for the period of ten years).

The greatest change in the administrative structure took place in the Faculty in 2010 when the Department of Education was split into 3 Departments: Department of Education and Special Education (*Institutionen för pedagogik och special pedagogic, IPS*), the Department of Education, Communication and Learning *(Institutionen för pedagogik, kommunikation och lärande, IPKL*), the Department of Pedagogical, Curricular and Professional Studies *(Institutionen för didaktik och pedagogisk profession, IDPP*) and the Department of Food and Nutrition, and Sport Science *(Institutionen för kost- och idrottsvetenskap, IKI).*

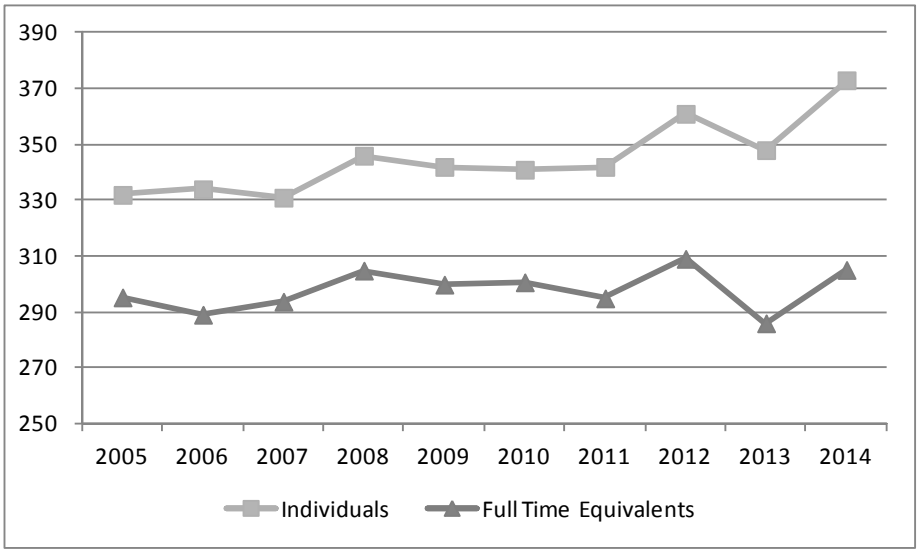

Figure 2 Change in the number of employees at the Faculty of Education (2005-2014)

The number of academic staff within the Faculty over the period of 10 years has remained relatively stable – it has ranged slightly from 295 in 2005 to 309 full time equivalents in 2014 (Figure 2). These statistical summaries were provided by Ulrika Nordberg Petersson, the Administrative director at the Faculty of Education, GU.

The Norwegian Model in GU

In GU, PRFS was introduced in 2008 to start in 2009, the same year as the national PRFS (University of Gothenburg & Universitetslednings kansli, 2008; *Prop. 2008/09:50*, 2008). The GU system

employs two kinds of performance indicators: bibliometric and the number of external grants (University of Gothenburg & Universitetslednings kansli, 2008, sec. 10). The proportion of the funding that is allocated on basis of bibliometric indicators varies from year to year. In 2009, the performance-based indicators were used to allocate 10 per cent of which bibliometric indicators accounted for half (5 per cent) (*Inriktningsbeslut för budgetarbetet 2012. Sammanfattning*, 2011). From 2012, this share has been raised to 15 per cent for the total performance-based share and 7.5 per cent for the bibliometric (Ibid.). Within the present system, the share of research funding that is calculated in relation to the total amount of funding for the year has reached 20 per cent (*Anslagsfördelning och kostnadsdebitering för budgetåret 2015 samt planeringsramar för 2016-2017*, 2014, p.4).

The data that are used in these calculations are collected in GUP that is affiliated to the GU library. As a representative from GU library explained in an e-mail (S.Gullstrand, personal communication, 12 December, 2014), GUP was set up in 2007 when it became *mandatory* to register all publications.

Every year in April there is a deadline for submitting data on publications for the purposes of annual research funding allocation. In rough lines, data on change in performance for a period of 4 years are transformed into a table with numbers representing what percentage share each of the faculties has in the overall figures for the whole university. Further, this share is translated into an additional research funding that is allocated to the Board of each of the faculties (see *Anslagsfördelning och kostnadsdebitering för budgetåret 2015 samt planeringsramar för 2016-2017*, 2014, p.4).

In addition, in 2011 a decision was made to adjust the distribution system for the funding in the Faculty of Education, the other two faculties of social sciences and the Faculty of Arts. It is claimed that due to insufficient funds within the faculties of social sciences, but spare funds in the Faculty of Arts, 30 per cent of the above-allocated funds will be redistributed on basis of the education 'volume' *(utbildningsvolym),* number of doctoral students and the same performance indicators – the number of external grants and publications. This is done in order to 'increase the overall competitiveness of the university' and it is expected that there will be results in 3 years time (*Inriktningsbeslut för budgetarbetet 2012. Sammanfattning*, 2011, p.2).

Since 2010 in GU, each faculty can freely choose between various kinds of bibliometric models that are used to distribute university funding among faculties (Henning, 2012). The Faculty of Education employs the Norwegian model. Here, I must note that a similar system is employed also on the Faculty level to calculate further distribution of funding among the Departments within the Faculty. This level of the use of PRFS is not explored in this study.

The Norwegian model is a model based on a differentiated approach to counts of *all* publications. Details will follow. This model was developed in Norway on a government commission for national research funding allocation, for a continuous monitoring of research processes on national and institutional level as well as to enable easily accessible data for various internal and external purposes such as construction of researcher's CV, institutional reports and so on (Sivertsen, 2010). These multiple goals are achieved by the construction of a data infrastructure that covers *all* publications, and in which these publication data are *transformed* into 'the indicator for the performance based funding model' (Sivertsen, 2010, p.23).

I have not found detailed information about the purpose of the implementation of the GU system, but from the available information it seems that this system is linked to the research evaluation project 'RED10' that was conducted in 2010 with a purpose to 'identify strengths and weakness in current and planned research' (Jarneving, Henning, & Sandberg, 2012, my translation). Pam Fredman, the vice-chancellor of GU suggests in the preface of RED10:

> *In order to deal with growing competition and the profiling requirements that are increasingly becoming a feature of higher education, we need to be able to maintain our position – not just nationally, but also globally.*
>
> (in Holmgren & Bertilsson Uleberg, 2011, p.5)

A similar formulation of purpose can be identified in GU Research and education strategy for 2009-2012 (Wallin Peterson, Weibull, & Elofsson, n.d.). Here it stated that the focus of this period is on 'the strengthening of the quality of research and long-term competitiveness' (Ibid., p.5, my translation). Competitiveness is named as a purpose also in a summary of a University Board meeting on annual budget for 2012 in 2011 (*Inriktningsbeslut för budgetarbetet 2012. Sammanfattning*, 2011).

The principle of $PRFS_{GU}$ is that the calculations of the performance of all the faculties are used to rank the faculties according to their relative performance. It is not the actual number in the performance indicators that determines the amount of the allocated funding; instead, it is the relative contribution of a faculty to the overall University performance. Thus, it is assumed that higher performance of a faculty is higher performance in relation to the performance of other faculties. However, it is not clear how appropriate such relational understanding of higher performance can be for educational research; therefore, this is one of the assumptions that I will explore in greater detail in relation to the quality of educational research.

The calculation for funding distribution purposes within GU is conducted annually, but the data are used from a period of the last four years split in two three-year periods. For example, a calculation for the purposes of year 2010 would use data on publications for the years 2005-2008 that are split in 2005-2007 and 2006-2008. In such a way the impact of annual fluctuations is decreased.

The conceptual implication of such a calculation is that it implicitly reinforces the time dimension in the publication process and in the academic work in general. Bibliometric performance becomes close to an assumption of *speed* in a production sense.

The Norwegian model used at the Faculty of Education at GU is adapted from the so-called Norwegian model that was developed by the Norwegian Association of Higher Education Institutions in 2004. The information on the Norwegian model I have acquired from scholarly texts that describe the original Norwegian model (e.g. Sivertsen, 2010; Schneider, 2009; Hicks, 2012).

The Norwegian model has a rigorous definition of what counts as publication and how the quality of data is achieved. In this model only *scientific* and *scholarly* publications are regarded as publications, and in order to be regarded as such 'a publication must

1. Present new insight
2. [be i]n a form that allows the research findings to be verified and/or used new research activity
3. [be i]n a language and with a distribution that makes the publication accessible for a relevant audience of researchers
4. [be i]n a publication channel (journal, series, book publisher) with peer review' (Sivertsen, 2010, p.24).

In these criteria there is an implicit assumption about peer review as warrant for quality. Schneider (2009) weakens such an assumption by suggesting that the fact that a publication has passed peer review renders it as 'an offer' to a research community, while the actual quality in a form of impact (citation count) or otherwise is yet to be seen.

An additional note is that in relation points 3 and 4, the publisher has to include authors from more than one institution: this criterion is added to exclude university-based publishers where there is a focus on publishing primarily the work of its own staff. Regarding the points 1 and 2, Sivertsen adds that these criteria have to be checked separately, since not all publications that are listed under a specific type or are published by a certain publisher, meet these criteria. These criteria are the next 4 criteria on what counts as a publication that have to be discussed in relation to the quality of educational research. For simplicity, I will split them in two separate assumptions: (a) consensus is possible in educational research community; (b) choice of the publishing language depends on the relevant audience.

Following the just-introduced criteria of a publication, a list of publishing channels is compiled. In this model, only those channels that has the International Standard Book Number (ISBN) and International Standard Serial Number (ISSN) are considered ('Procedures for processing new submissions', n.d.). The responsible administrative unit for the inclusion of new channels is The National Publishing Board ('Forslag til vedtak av nye kanaler på nivå 1', 2016). In addition, GU bibliometric unit can include a publishing channel, if it meets the above-describe criteria – it has ISBN/ISSN number and peer-review is employed ('Gothenburg University Library: Allocation', 2015).

Following this definition of publications, the incentive is to include all publications. To ensure that data on publications are of high quality, Sivertsen distinguishes between four levels of data recording and control: individual, departmental, faculty/university and national level. He emphasises the need of co-ordination, communication and continuous transparency among all the levels. The full inclusion of publications is taken into account also in the GU system – the data for the Norwegian model are acquired from a GU publication database (GUP). All the data are reported by researchers themselves. In GU, the quality of the data is assumed to be the responsibility of each Department (B. Jarneving, personal communication, April 20, 2016).. However, the data are checked by GU librarians taking the corresponding actual publication as the reference for the entry in the database (Ibid.). Thus, the quality assurance in GU corresponds to the one that is assumed to be in place for the original version of the Norwegian model. This reliance on the checked quality of data as determining the accuracy of calculations could be another assumption to explore; yet, it will be left beyond the scope of this study.

The next step is the affiliation of authors and institutions. In the original Norwegian model, all the authors are indicated at the data entry process. Hence, the publication can be attributed not only to the first author, but to all who had been listed as authors. In addition, it is possible to conduct further analysis employing fractioned counting. In GU, the authors and the institutions are reported by the researcher who submits the data, but, as mentioned above, all the bibliographic data are checked by librarians.

To assure comparable measurement, the model employs *weighting* of different *types* and different *levels of prestige* of publications. The model includes four types of publications: authored (not edited) books, book chapters, peer-reviewed journal articles and conference proceedings. For conference proceedings, if is it is published in a channel with an ISBN number, then it is regarded as a book chapter. If it published in channel with an ISSN number, then it is regarded as peer-reviewed journal article.

In addition, the included publishing channels are evaluated by the National Publishing Board whereby those channels that are assumed 'the "most selective" and "prestigious" channels' (Schneider, 2009, p.371) are assigned to a separate higher 'level'. Consequently, each publication included in the model is weighted to points of the corresponding level (Table 1 Publication points in the Norwegian model).

Table 1 Publication points in the Norwegian model

	Peer-reviewed journal article	Book chapter	Book
Level I	1	0.7	5
Level II	3	1	8

(Sivertsen, 2010; 'Gothenburg University Library: Allocation', 2015)

Schneider (2009) interprets this weighting as an attempt to go 'beyond mere productivity' (p.365), and on the same lines, Sivertsen (2010) confirms that the intention behind such a differentiation is that the

> *...high "Level 2" is meant to stimulate the institutions to encourage their researchers to seek the most competent and critical peer review and a wide relevant readership. We also want to avoid a tendency to increase publishing in less significant channels.*
>
> (Sivertsen, 2010, p.26)

However, Sivertsen (2010) acknowledges that this division in levels is 'difficult and controversial'. Therefore, the set of the Level 2 channels is revised annually by the National Publishing Board in collaboration with National councils of each discipline. Such an approach is assumed to be more reliable for SSH where the typical citation-based indicators such as journal impact factors might be misleading.

To maintain the level of difference between the higher level publishing channels and the others included in the evaluation model, criteria of no more than 20 per cent of world publications of the field can be assigned to the higher level (Sivertsen, 2010; 'Gothenburg University Library: Allocation', 2015). Earlier I mentioned that the GU bibliometric unit can include other channels that might be assumed relevant within GU, but are not included in the Norwegian list. This can be done only with respect to the Level 1 publications ('Gothenburg University Library: Allocation', 2015). The Level 2 list is kept as in the original list.

The present 'Norwegian' list of approved journals in the category 'Pedagogy and Education' contains 925 titles of whom 744 are assigned to Level 1, while Level 2 contains 73 titles ('The Norwegian Register for Scientific Journals, Series and Publishers: Scientific journals, series', 2016). In comparison, the present list of educational research journals (WoS categories: 'Education & Educational Research'; 'Education, special') in WoS Social Science Citation index (SSCI) contains 272 journals ('SOCIAL SCIENCES CITATION INDEX - EDUCATION & EDUCATIONAL RESEARCH - JOURNAL LIST', 2016; 'SOCIAL SCIENCES CITATION INDEX - EDUCATION, SPECIAL - JOURNAL LIST', 2016).

Hicks (2012) and Schneider (2009) suggest that in setting up an independent unit that represents the interests of particular research fields in determining what qualifies as relevant publishing channels, this model addresses the core problem of using databases such as WoS where only a small part of publications are included in the database. However, precisely this point has received critique from the Swedish community. Henning (2012) and Hammarfelt, Nelhans, Eklund, & Åström (2016) have suggested that the Norwegian model is accused of being *Norwegian*. Potentially, there is a risk that some publishing channels that might be considered of value in GU but not in Norway are not included in the list. This problem is addressed in GU in the following way. The GU bibliometric unit can include a publishing channel, if it meets the criteria of a publication that are used in the Norwegian model ('Gothenburg University Library: Allocation', 2015). Alternative to this would be a separate national list of journals as it is done in Denmark (on the plausibility of this see Henning, 2013).

However, an argument against could also be put forward, as has been done by Sandström and Sanström (2009), who favour WoS data also in SSH. Their argumentation justifies WoS since it provides high quality data. In a way, they suggest that there are good reasons why the majority of SSH publishing channels are not included. Under their proposals the choice of publishing channels would not be dependent on the sense of quality of one particular national research community. This is where it is no longer clear which research community – the Norwegian (behind the *Norwegian list*) or the international (behind the WoS list) has a more trustworthy and appropriate sense of quality for educational research as it is seen, for example, within GU.

The final step that is conducted in the calculation of the performance indicators is fractioning the publication counts. In the Norwegian model, the publication counts are fractioned by the number of authors and the number of institutions. This means, for example, that if a publication has two authors and one of whom is employed in two institutions, the final number of fractions that is attributed to a particular institution is derived from the scheme showed in Figure 3 on page 38. In GU, the fractioning is done in the same way (B. Jarneving, personal communication, April 20, 2016).

However, in the original Norwegian model, if a publication has more than 10 authors, an institution is still assigned a minimum of 1/10 of a publication point (Sivertsen, 2010), which is the practice also in GU. A point to be made is that Schneider (2009) suggests that the current practice in Norway does not follow this principle, but he does not describe the existing practice. The meaningfulness of this fractioned counting could be another assumption to explore, yet also this one will be kept beyond the scope of study.

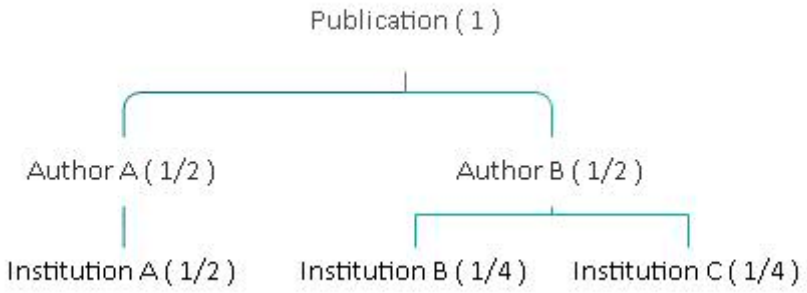

Figure 3 Scheme of publication fractions

As the final comment about the Norwegian model, I would like to alert to the fact that Sivertsen (2010) does explicitly say that these publication points are meant to be used only for the purposes of funding allocation on the level of institutions. It is not meant to be employed by various external institutions, such as research councils, or to use it to measure performance on individual level.

Regarding undesirable consequences from the use of the Norwegian model, Schneider (2009) comments that adverse effects of the Norwegian model are not known. A potential undesirable effect could be increase merely in the number of publications – increase in the *quantity* as it has occurred in Australia (Butler, 2003). Sivertsen (2010) claims that the Norwegian model has avoided the risk of merely promoting the quantity of publications by introducing the Level 2 research publications.

Possibly, at that time – more than 5 years ago - it was too early to identify potential adverse effects, but just recently, Hammarfelt and de Rijcke (2015) made such an attempt within humanities at the Uppsala University, as I reported earlier. As a reminder, their conclusion was that there is, for example, increase in the number of publications in English and among scholars concerns have been expressed on the changing sense of the worth of the various publication types that can be found in the publishing traditions that are beyond the peer-reviewed journal article-dominated fields. Yet, the authors point out that it is not possible to claim that these changes are associated precisely with the use of the model. There may be other factors contributing to the shift in publishing practices.

Assumptions of the Norwegian Model

In this section, I acquire the final 'tool' that is required in order to interpret the publishing patterns that I aim to explore and report in the next section. In the description of the model I have made several notes on the assumptions on which the model is based.

Following the earlier introduced insights from social epistemology, I identify *goals of research* that are implicit in the design of the Norwegian model and its use in GU. Here, I introduce these assumptions once more, but this time already combining them into sub-themes: (1) the role of consensus, (2) of language and (3) of bibliometric performance in research. Given the focus of this study, I reformulate these assumptions in relation to educational research.

The main assumption about the goals of research can be identified in the 'object' – bibliometric performance - that Norwegian model measures and the consequences that follow from this

measurement. In the Norwegian model, higher bibliometric performance is seen as higher performance of a faculty relative to the performance of other faculties.

If such an assumption is true (in a logical sense), then for the research question in which I attempt to identify compliance in publishing patterns in GU educational research (2005-2014) with assumptions implicit in $PRFS_{GU}$, it becomes desirable to see change in publishing patterns in the following terms:

(1) Increase in the number (quantity) of Level 2 publications that score higher
(2) If (1) is identified already before the introduction of PRFS, then increase in the average annual growth rate (speed) of Level 2 publications that score higher.

Given the fact that in GU the Norwegian model is used by calculating the sum of all publications, whereby the difference of Level 1 and Level 2 is lost, it could be that it is desirable to see any increase in quantity or speed for any publications that are assigned with publishing points. Table 2 shows what such reasoning might look like in relation to types of publications.

Table 2 Relative worth of publications

Type	Points
Book, Level 2	8
Book, Level 1	5
Peer-reviewed journal article, Level 2	3
Book chapter, Level 2 and Peer-reviewed journal article, Level 1	1
Book chapter, Level 1	0.7

For peer-reviewed journal articles, an additional preference can be identified in relation to language. As I described earlier, all the journals that are assigned to Level 2 publish only in English. Thus, implicitly a priority is given to articles written in English, on a condition that they are published in Level 2 journals.

In addition to bibliometric performance as a goal, the Norwegian model implicitly assumes that there is consensus in research community on various aspects of research. This assumption can be identified in the very essence of the Norwegian model, namely the very definition of publications, whereby it is implied that consensus is possible in educational research community

(a) On quality of research
(b) On what is 'new insight'
(c) On how research findings can be 'verified and/or used in new research activity'
(d) On which journals or publishers should be regarded as 'more prestigious'.

Given the great diversity of educational research, it may be the case that, even if it may be plausible to aim for higher bibliometric performance in the sense just introduced, the Norwegian model in the context of educational research is not valid. This possibility I will discuss along with the assumption about bibliometric performance in the chapter on discussion.

As a final note I can add that the above introduced list of assumptions is not exhaustive. Several other assumptions could be identified, such as the meaningfulness of fractional counting of publications, the reliance on data infrastructures and numbers, the role of time in bibliometric performance, and more. Not exploring these further assumptions might be regarded as a limitation of this study, yet I suggest that the assumptions I do explore in this study provide enough material to argue for two opposing positions towards the use of PRFS in educational research.

Findings

In this chapter, I present the findings from my empirical exploration of change in publishing patterns in educational research (2005-2014) before and after the introduction of the GU performance-based research funding system (PRFS) in 2009. Here I provide my answers to the questions that I have earlier set out:

- What change (if any) can be identified in the number of peer-reviewed journal articles in relation to:
a. The categories used at the funding system and language?
b. The categories used at the funding system and other types of publications?
- Is there compliance in publishing patterns in GU educational research (2005-2014) with assumptions about research implicit in the GU performance-based research funding system introduced in 2009?

In my description of findings I follow a principle from general to specific: I begin with general figures on the total volume of publications in relation to the PRFS and publication type. After that, I describe patterns of change I have identified in relation to the introduction of PRFS in 2009. First, I report general figures of change in the total volume of publication. Then, I continue with general figures of change for other types of publications. In the description of articles, I provide a detailed analysis of peer-reviewed articles and change also in the use of language for this type of publications. Finally, I summarise these findings in relation to the research question. Descriptive statistics can be found in Appendix 5 and Appendix 6.

In this dataset, I explored 4090 publications distributed over a period of 10 years. In relation to the publication type, the largest share (21 per cent, n=852) of the publications are book chapters. The next largest is peer-reviewed journal articles, which accounts for 18 per cent (n=753) of the total volume of publications. Books together with textbooks account for 4 per cent (n=151). About a third of publications (31 per cent) are conference proceedings and other conference contributions. A relatively large share of publications in this data can be identified in the categories ‘other articles’ (15 per cent) and reports (5 per cent). On details for descriptions of these and other types of publications see Appendix 1.

Exploration of change in the total volume of publications suggests that throughout the period of 10 years there has been a steady increase in the number of publications. There is a minor (1 per cent) difference between the average annual growth rate before the introduction of PRFS (59 per cent) and after (60 per cent). The annual number of publications has doubled from 274 in 2005 to 527 publications in 2014.

The change in the share of publications included in PRFS, similarly indicate a steady increase over the whole period, with a drop in 2010. The annual average growth for the number of publications that are assigned either to the Level 1 or to the Level 2 has changed from 65 to 69 per cent. It might be that the drop in Level 1 articles in 2010 is linked to the restructuring of the Faculty of Education. For example, the time that could have been spent on writing publications in 2009 might perhaps have been spent in uncertainty, or indeed on preparatory work for change in the faculty structure.

Exploring the change in relation to publication type, I would expect that the increase in the number of publications will be for only those publication types that are included in PRFS. The data however suggest that this is not the case. For almost all types of publications the share of publications in 2010-2014 is greater than in 2005-2009. This means that there has been increase in the number of publications already before the introduction of the PRFS in 2009.

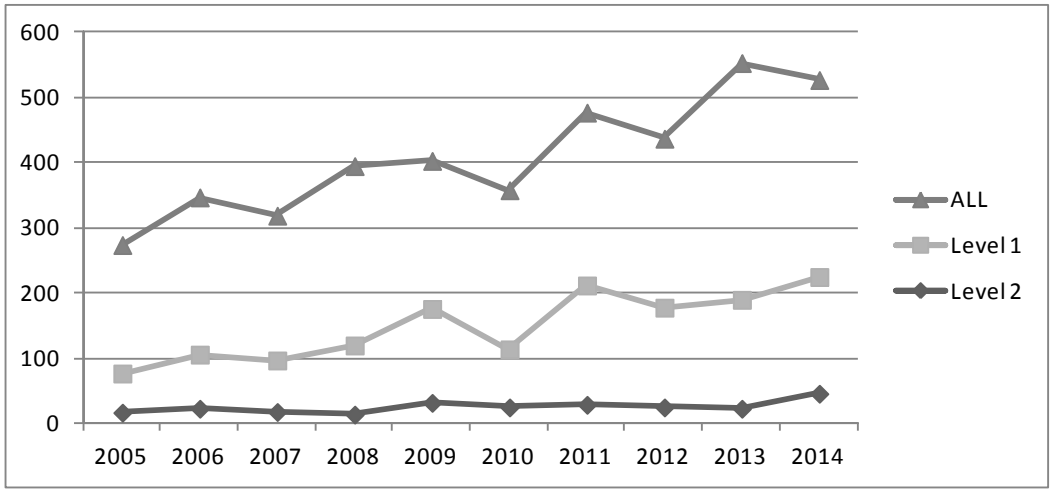

Figure 4 Change in the number of publications, by PRFS categories

The exceptions for this pattern are reports and 'conference paper peer-reviewed'. The percentage change for these types is -71 and -6 per cent respectively. The reason for the change in the number of these conference papers maybe a misunderstanding of the categories of publications that is used in GUP. The pattern in Figure 5 (page 41) may indicate that before 2012 conference *abstracts* have been submitted to the system as conference *papers.* This is merely a suggestion without any supporting evidence. Therefore, further interpretation of this publication type will not be conducted.

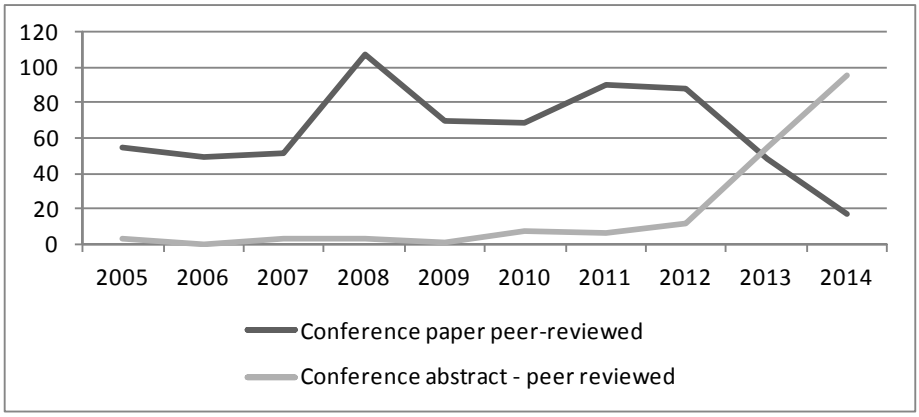

Figure 5 Change in the number of conference papers

Further inquiry into the change over time in the numbers of publications indicates that there is a rather great difference in the patterns of change among the various publication types (Figure 6). A considerable increase can be observed for peer-reviewed journal articles. The average annual growth rate in 2005-2009 was 58 per cent, but in 2010-2014 it has risen to 65 per cent. In counts, it means that from 40 articles in 2005 the number has risen to 60 in 2009, but then from 68 articles in 2010 the number has changed to 137 peer-reviewed journal articles in 2014. I will return to further details in changes in peer-reviewed journal articles in relation to PRFS categories and language shortly, but first I continue with figures on other types of publications.

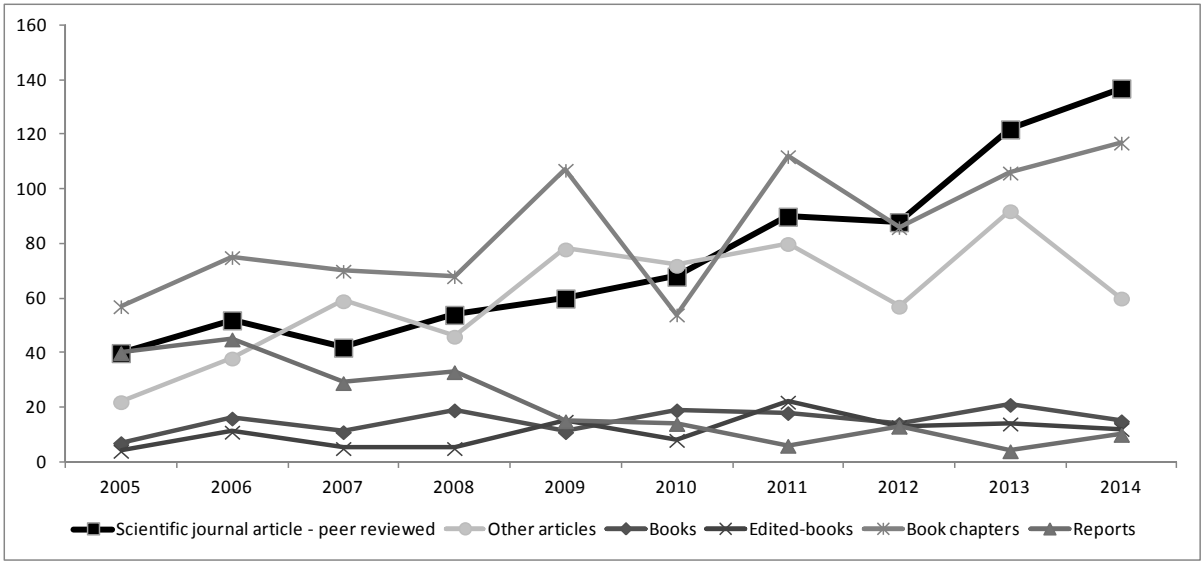

Figure 6 Change in the number of publications, by type

An increase over time can be observed also for book chapters and 'other articles', which includes non-peer reviewed scholarly articles, review articles, book reviews and 'other articles' (this is the original category within GUP). From this group of publications, only 4 per cent (n=22) are included in PRFS.

For book chapters, the average annual growth rate is quite considerable: during the four years between 2005 and 2009, it has been 60 per cent, but in the period between 2010 and 2014 it has risen to 71 per cent. A noteworthy pattern for book chapters can be observed in relation to PRFS categories (Table 3). The annual average growth rate has increased considerably (by 27 per cent) for those book chapters that are *excluded* from PRFS.

Table 3 Change in the number of book chapters

	Full counts			Percentage rate	Annual average growth	
PRFS	All	2005-2009	2010-2014	PR	2005-2009	2010-2014
All	852	377	475	26%	60%	72%
All included	565	257	308	20%	68%	71%
Level 1	484	223	261	17%	70%	72%
Level 2	81	34	47	38%	61%	67%
Excluded	287	120	167	39%	48%	75%

The number of books and edited books seem relatively stable over the whole period of 10 years. For books the annual average growth rate in 2005-2009 was 75 per cent, but in 2010-2014 it has decreased to 49 per cent. In relation to PRFS Levels, in the period before the introduction of PRFS there was only one Level 2 book, but after (2010-2014) there were 5 books assigned to this *higher* level. The number of Level 1 books has increased by 2 books (from n=40 to n=42).

For edited books, the annual average growth rate has decreased from 78 to 71 per cent in 2010-2014. This means that there has been a greater increase in the number before the introduction of PRFS, but after the introduction, the growth has diminished. Similar patterns of diminished average annual growth rate can be observed also for textbooks and doctoral theses. Descriptive statistics as well as growth rates for conference papers, theses and text books are provided in Appendix 5.

Now, I have a closer look at the number of peer-reviewed journal articles in relation to both the PRFS categories and the language of the article (Figure 7 and Figure 8 on page 43). Almost all (94 per cent) of peer-reviewed journal articles are included in PRFS (assigned either to Level 1 or Level 2). For Level 1 the share is 72 per cent of all articles of this type, but for Level 2 it is 22 per cent. In relation to language, the largest share (87 per cent; n=655) of peer-reviewed journal articles are articles written in English. Articles in Swedish account only for 11 per cent (n=86), and articles in other languages for only 2 per cent (n=12). A cross-tabulation of the number of peer-reviewed journal articles with PRFS categories (Level 1 and Level 2) and language indicates that all but two of Level 2 articles are in English, the other two are in Swedish.

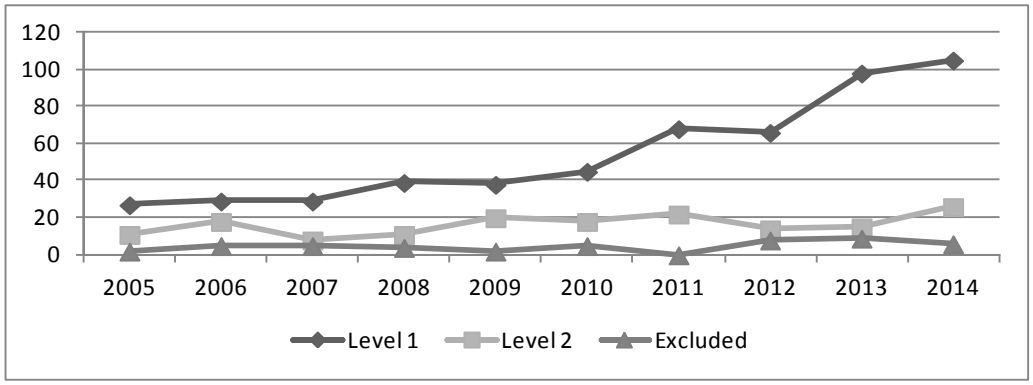

Figure 7 Change in the number of peer-reviewed journal articles, by PRFS categories

An exploration of the patterns of change suggests that before the introduction of PRFS, the average annual growth rate for all these articles in English was 61 per cent, but after the introduction it has risen to 68 per cent. In relation to levels, the growth rates are diverse: the patterns indicate that after the introduction of PRFS there is an obvious (20 per cent) increase in Level 1 (from 58 to 78 percent) peer-reviewed journal articles in English. This is not the case for Level 2 articles - growth rate has decreased by 10 per cent (from 61 to 51 per cent). The average annual growth rate for peer-reviewed journal articles in Swedish has increased only by 1 per cent (from 47 to 48 per cent).

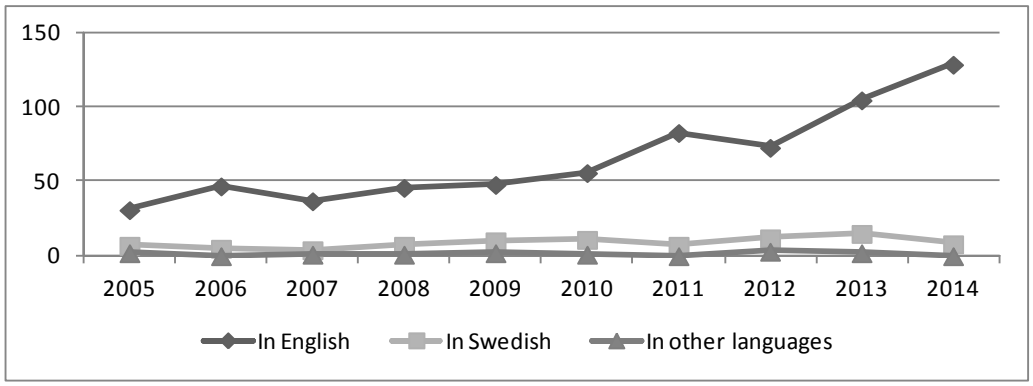

Figure 8 Change in the number of peer-reviewed journal articles, by language

In summary, in this empirical exploration of publishing patterns in educational research (2005-2015) in GU I have attempted to identify and describe change in relation to the introduction of the GU PRFS in 2009. The focus of this study was compliance in publishing patterns in GU educational research (2005-2014) with assumptions about research quality implicit in the GU performance-based research funding system introduced in 2009. Therefore, I have opted for exploring one of the prioritised types – peer-reviewed journal articles in English. The research question that captured this focus was 'What change (if any) can be identified in the number of peer-reviewed journal articles in relation to: (a) the categories used at the funding system and language and (b) the categories used at the funding system and other types of publications?'

The findings of this study suggest that, firstly, there is change in the number of peer-review journal articles. For the period of 5 years before the introduction of PRFS (2005-2009) there 248 peer-reviewed journal articles, but in the 5-years after the introduction (2010-2014) this numbers has almost doubled (n=505). A change in the number of publications was identified for all the types of publications, but the patterns differ from those identified for the peer-reviewed journal articles. For almost all publications, the total number of publications in the 5-year periods after (2010-2014) the introduction of PRFS is greater than in the period before (2005-2009). The only exception is reports where a considerable decrease was observed. In contrast, the increase in the average annual growth rates was identified only for peer-reviewed journal articles and book chapters. For other types the annual average growth either has remained the same or decreased.

In relation to the PRFS categories, the only considerable increase in the number of publications that are included in PRFS was for peer-reviewed journal articles. The annual average growth rate differences suggest that such an increase in growth rate is greater for Level 1 than for Level 2. For book chapters, a greater increase can be observed in the number of those publications that are excluded from PRFS. A slightly different pattern can be observed Level 2 books. For this publishing category there has been change from only 1 book in the period from 2005 to 2009 to 5 books in the period from 2010 to 2014. Finally, a point of reminder is that the above reported finding of increase in the total 5-year period number of almost all publications (exception: reports and conference papers) refers also to those types of publications that *are excluded* from PRFS.

In relation to language and PRFS categories, my findings suggest that there is more rapid increase in the Level 1 peer-reviewed journal articles in English after the introduction of PRFS, than before. In contrast, for Level 2 articles in English there is an increase, but the growth rate after the introduction of PRFS has slightly diminished. For Swedish, the average annual growth rates are relatively stable: for the number of articles in Swedish that are included in PRFS the growth has remained the same (45 per cent), but for Level 1 it has slightly (5 per cent) decreased (from 50 to 45).

In relation to the research question on compliance, it seems that the change in publishing patterns does not exactly match the categories prioritised in PRFS, thus indicating that there is not wholesale compliance with the implicit assumptions of the PRFS. If assumptions are seen in three broad categories – quantity, speed and relative worth of publications – then it seems that quantity has been a priority already before the introduction of PRFS. This is indicated by the fact that there is increase in all (but reports) publications after the introduction of PRFS. In relation to speed, this assumption may be clearly identified only for peer-reviewed journal articles and book chapters. In relation to the relative worth of publications, then here too only the change in the average annual growth for peer-reviewed journal articles suggests compliance with an assumption that this is a desirable publishing channel. However, such an increase is only for the lower level (Level 1) articles, but not for Level 2 articles. The only other change that may indicate compliance with the assumption of relative worth is the change in the annual average growth rates for Level 2 books. Given the fact that these increases capture only a few patterns among those that may be seen as prioritised within PRFS, my conclusion is that the findings of this study suggest that the change in the publishing patterns does not indicate full compliance with the implicit assumptions in PRFS.

Before I continue with a discussion, I wish to acknowledge a few limitations. While exploring the data I discovered that 19 publications were affiliated to an institution that in the year of publication either no longer exists or is not yet established, thus indicating a potential source of error in the data. However, due to the fact that this number amounts to less than 0.001 per cent of the total number of publications, these entries were kept in the dataset without adjustments. Another limitation was the fact that my exploration of publishing patterns for such a relatively small dataset, but with a great number of categories that are cross-tabulated, resulted in zero values for some types of publications. Because of these zero values it was not possible to calculate growth rate for those publication types and hence compare all the publication types in all the sub-categories.

Discussion

In this chapter I discuss the findings of the study in relation to the literature review that I have presented earlier. Due to the various levels of abstraction in my theoretical lens of social epistemology, I discuss the findings accordingly. I begin with the empirical level – the components of the total volume of publications and the changes in publishing patterns in relation to PRFS. Then, I discuss these changes in publishing patterns in relation to the quality of educational research. In the final part of the discussion, I reflect upon how the choice of social epistemology as a theoretical framework aided or impeded the conceptualisation of implications from PRFS to the quality of educational research. After this, I explicate some general key insights from this study and note a few further directions for research on the same topic.

Publishing Patterns within Educational Research

In this study I explored publishing patterns within the field of educational research at the University of Gothenburg (GU), in Sweden. The time frame of this study was a period of ten years (2005-2014) in the middle of which a PRFS was introduced both on the national and university level. The aim of this empirical part of the study was to understand how the publishing patterns within the field of educational research GU changed in relation to the introduction and use of the institutional performance-based research funding system.

I explored this dataset (n=4090) on publishing patterns in relation to 16 publication types. Since in most other studies the only types that are explored are peer-reviewed journal articles, book chapters and books, in my discussion too I summarise results in relation to only these types (see Table 4 on page 45).

Table 4 shows that findings of this study in relation to the distribution of publications across types with minor differences resembles the average pattern for 3 Swedish universities during the period of ten years before the introduction of PRFS in 2009 (2000-2009) (Hansen & Lindblad, 2010). This would suggest that a major shift in the publishing culture within Swedish educational research at GU has not happened after the introduction of PRFS.

Table 4 Comparison of publishing patterns within the field of educational research

	Sweden [1] (this study)	Sweden [12a]	Norway[3b]	Flanders [3b]	Germany[1c]
	2005-2014	2000-2009	2005-2009	2000-2009	2004-2006
Peer-reviewed journal articles, %(n)	43 (753)	43 (920)	44 (1102)	92 (1156)	35 (1568)
Book chapters, %(n)	49 (852)	47 (1000)	47 (1174)	5 (63)	49 (2192)
Books, %(n)	7 (130)	9 (200)	9 (216)	0.05 (0.63)	16 (695)
Total, N	1735	2120	2492	1220	4455

1 sample of national data

2 figures are approximate

3 fractionalised counting

Sources: a-Hansen & Lindblad, 2010; b-Sivertsen & Larsen, 2012; c-Engels et al., 2012; d-Dees, 2008

Comparing these patterns to those within other countries, a few considerations have to be reminded. Earlier I noted that the cross-country comparisons that are based on diverse bibliometric studies may be misleading. To decrease this risk, I discuss the patterns using merely proportions of publications in relation to type and language assuming that even if the numbers are not exactly comparable, they may indicate to an approximate sense of national publishing patterns.

Patterns identified in this study are close to the ones found in the total volume of Norwegian publishing patterns (Sivertsen & Larsen, 2012). In contrast, the findings of this study differ quite

substantially from the ones in Flanders (Engels et al., 2012). In Flanders, the largest part of the publications are peer-reviewed journal articles, with only a minor part (less than 6 per cent) of books and book chapters. However, this comparison to Flanders and also to Norway might be misleading, since these two studies employ fractionalised counting of publications.

Findings from this study also differ from those in Germany (Dees, 2008). In contrast to GU, in German educational research there seems to be much greater share of publishing in books (16 per cent in contrast to 7 in GU) and less publishing in peer-reviewed journals (35 per cent in contrast to 43).

Within this study, the exploration of the use of language within peer-reviewed journal articles suggests that in 87 per cent of all articles of this type are published in English. This share varies across different countries. In Norway, the number varies from 33 per cent (Sivertsen & Larsen, 2012) to 47.2 per cent (Ossenblok et al., 2012), but in Flanders, it is 85 per cent (Ossenblok et al., 2012). From Germany, data specifically on the language for peer-reviewed journal articles are not given, but it is suggested that on average the share of English language was about 11 per cent (Dees, 2008).

In sum, this comparison suggests that in publishing patterns within the field of educational research some similarities can be identified among the reviewed countries (the dominance of book chapters). However, the case of Flanders with the dominance of peer-reviewed journal articles indicates that it might be the case that variation across countries is much greater. In contrast, the variation in relation to the language used when writing peer-reviewed journal articles is much greater. Here, one can identify equally used English and national language (this study, Sweden - Hansen & Linbdlad, Norway - Ossenblok et al., 2012), dominance of English (Flanders - Ossenblok et al., 2012) or national language (Germany - Dees, 2008).

Exploration of compliance with assumptions of PRFS $_{GU}$, showed that compliance with all the assumptions of $PRFS_{GU}$ was not identifiable in this study. Yet, some patterns of compliance are distinguishable if change in publishing is explored in relation to each of the implicit priorities, such as quantity and relative worth of publications. As a reminder, it was possible to identify that for all (but reports) the number of publications *after* the introduction of $PRFS_{GU}$ in 2009 was greater than before. This means that already before 2009 there has been emphasis on *quantity* and this emphasis has remained also in the presence of $PRFS_{GU}$. It may be the case in GU there has been already some other incentives apart from PRFS that prioritise publishing. It might be that it was known already in 2005 (the first year covered in this study) that in few years bibliometric indicators will be introduced both on national and institutional level. Available information does allow me to suggest that this is the case.

If it is the case that the changes I have identified in this study can be explained with other factors, it does not change the fact that there is PRFS in GU, and if there is PRFS, then it is meant to 'work' in one way or another, despite the fact that it is highly problematic to provide empirical evidence that would link the changes in publishing patterns and the use of PRFS (see Hammarfelt & de Rijcke, 2015; Osuna et al., 2011). The aim of this study was not the establishment of causality between PRFS and change in publishing patterns, but the awareness of this epistemic complexity adds another question: what PRFS are actually aiming to achieve if the goals that are assumed to be the goals cannot be shown empirically? With this question in mind I will discuss the change in publishing patterns in a more detailed manner.

If the other patterns of potential compliance are rendered through the lens of the Norwegian model, in which the Level 2 is introduced in order to differentiate a 'more prestigious' publishing channels, then the findings of this study may indicate that for educational research there is an increasing move towards *quantity*, but not *quality* (in the sense of the Norwegian model). The same comment would apply also for the continued publishing in a form of book chapters, which is a channel that in the Norwegian model is rendered *less* worthy than peer-reviewed journal articles. An exception would be the change for Level 2 books. This may be the only indicator in these data that, according to Norwegian model, indicate a change *towards quality*.

Leaving for a moment the Level 2 books as an exception, the patterns for peer-reviewed journal articles and book chapters reflect findings from Australia (Butler, 2003) where introduction of PRFS manifested in publishing patterns as merely enhancing *quantity*. Butler suggests that a differentiation in terms of quality of the various publications might be a solution. This comment was taken into consideration in the design of the Norwegian model (Sivertsen, 2010) and hence the Level 2 was introduced to promote publishing at this *higher* level. An empirical exploration of publishing patterns in Norway, indeed suggests that the aim seems to be achieved – higher increase is observable for Level 2 publications (Sivertsen, 2010). Thus, findings from this study of educational research publishing patterns in GU in a context of the GU_{PRFS} seem to contradict the incentives built in to the design of the Norwegian model.

Findings of this study also differ from those conducted recently in the Faculty of Arts in Uppsala University (Hammarfelt & de Rijcke, 2015). On the level of publication types, patterns from Uppsala (2006-2013) do not indicate a particular pattern of change, but the patterns in GU indicate an apparent increase in the number of peer-reviewed journal articles and also decrease in reports.

Another pattern that can be compared is change on the level of language. In this study, I explored change in the use of language only for peer-reviewed journal articles and I found that increase can be observed for both languages, but it is higher for English. In combination with levels, my findings suggest that for Swedish articles, the change in the average annual growth rate is minor. In the study on patterns from Uppsala, it is reported that there is decrease in publishing articles in Swedish, but increase in publishing them in English. However, a noteworthy point from the study by Hammarfelt and de Rijcke (2015) is that it is not possible to conclude that these patterns correspond with the assumptions implicit in PRFS – it cannot be concluded from the data alone that the change is indeed due to the introduction in PRFS.

These main points in the literature – including the manifestation of incentives of the Norwegian model reported by Sivertsen (2010) – seem to indicate that the Norwegian model does not seem to *work* as it is *meant* to work in educational research in GU. Yet, the comments from Hammarfelt & de Rijcke (2015) alert to the fact there may be various explanations for change in publishing patterns, thus it is not easy to conclude that change that has been found, has to do with the incentives of the model. This is not surprising, since after all the problems with claiming causality between PRFS and change in publishing patterns were already considered in the design of the study.

Now, it is time to render these findings through the lens of social epistemology. By doing so, I distance myself from the reasons of change in publishing patterns. If I would take the Norwegian model as a legitimate way to transform the worth of educational research into numbers for research funding distribution purposes, then it would be plausible to say that the changes in the publishing patterns of educational research in GU does not indicate enhanced quality. This is the claim that I take as the starting point for the final section of this thesis – a theoretical discussion of implications from the use of PRFS on the quality of educational research.

Through the Lens of Social Epistemology

In this section of the discussion I provide an answer to the research question on the likely implications that the introduction of the GU performance-based research funding system has on the quality of educational research.

In order to identify the likely implications, I begin with the main assumption within social epistemology – the assumption that *knowledge is social*. In the chapter on social epistemology, I argued that if this assumption is transferred to the question about the quality of educational research, then the conclusion can be reached that knowledge about the quality of educational research is itself also social. Let us consider how such an assumption refers to bibliometric indicators that are used in $PRFS_{GU}$, what implications follow for the quality of educational research from the introduction of $PRFS_{GU}$ and, finally, whether such implications seem relevant to a context of other kinds of PRFS.

Understanding bibliometric indicators in the sense of social epistemology requires reconsideration of what kinds of knowledge claims are made when employing bibliometric indicators in research funding systems. In the chapter on bibliometrics, I referred to a wide-spread understanding that bibliometric indicators are measures of performance which is not the same as quality, but may be used to acquire an approximate sense of it, especially when some sort of differentiation between quantity and quality is introduced (Sivertsen, 2010; Sandström & Sandström, 2009).

This reasoning somewhat resembles Locke's distinction between primary and secondary qualities, which I referred to in the chapter on social epistemology. As a reminder, according to Locke primary qualities are those that are intrinsic to the 'body', while secondary qualities are those 'which in truth are nothing in the objects themselves but power to produce various sensations in us by their primary qualities' (Locke, 1990, p.90). Applying this to bibliometric indicators, the 'body' might be likened to educational research, so that the *primary qualities* become the quality of educational research in an epistemic/social sense, while publications are the secondary qualities of educational research which 'produce various sensations in us by their primary qualities'. Which means: if educational research is of quality, then it will also be manifest in the publications – the material aspect of educational research. Hence publications can be used to acquire some sense of the quality of research. This seems like a plausible argument. Yet, in the chapter on theoretical framework, I argued that already in the 18th century it was shown that such an argument is false (e.g. Kant, 2004; Hegel, 2013).

The false distinction between primary and secondary qualities of research is precisely what is also acknowledged in social epistemology (Fuller, 2002). If I treat publications as a representation of quality that is intrinsic to educational research, I neglect the fact that each of these concepts – quality (in epistemic or social sense) and research – has itself a social origin. It means that the meaning of these concepts is embedded in contexts of usage. The 'existence' of these concepts is dependent upon people who meaningfully act in relation to whatever may be referred to. Hence, if bibliometrics are worthy of consideration as markers of research quality, then their use is plausible only if in the corresponding research community there is consensus about the nature of the relationship between publications and the quality of educational research.

Is such consensus possible in the field of educational research? Let us recall some characteristics of this research field. One of its main characteristics is the diversity in conceptual sources and in research agendas that spans national and even institutional contexts. Consequently, there is great diversity also in conceptualising the quality of educational research. In addition, educational research has been criticised for being of low quality, irrelevant, lacking rigour (Oancea, 2005), but it is not easy to understand whether the reason for these critical comments has been due to the incommensurability of the various research paradigms or there are indeed grounds to doubt the strength of epistemic warrants for knowledge claims put forward by educational researchers. This diversity in itself does not suggest that consensus is not possible or indeed desirable. It is however clear that such a dialogical, unstable notion of consensus is one very different from the notion of consensus discussed in relation to the $PRFS_{GU}$ – The Norwegian model.

The core of the Norwegian model is the definition of 'publication'. Using the Norwegian model, bibliometric performance is analysed by means of publication counts. One of the criteria that a publication has to meet in order to *count* as a publication, is peer review. If peer review is seen as a filter whereby a piece of text is assigned the status of *an offer* to a research community (Schneider, 2009), there has to be consensus on how such an offer can be identified. To make it clearer, I will use Kuhn's (1970) incommensurability thesis. Given the plurality within educational research, it is hard to image that there can be achieved *cross-paradigm* consensus on criteria for what counts as 'an offer' to a research community. See for example the article that calls for more randomised-control-trials in educational research (Torgerson & Torgerson, 2001) and a harsh critique of evidence-based movement where randomised-control-trials are seen as the most reliable evidence (Welch, 2015). How likely would it be that scholars taking such hard lines on method would reach consensus on what should count as 'an offer' to the research community? Would they even agree on what constitutes the research community? My response would be – unlikely.

Further, in the Norwegian model, the inclusion of publishing channels is decided by research community and there is differentiation of publishing channels in two levels – no more than 20 per cent of those publications that are considered more worthy by research community are assigned to 'the higher' Level 2. Recalling the critique that has addressed the scientific worth of educational research, it is possible that by including all the channels that are proposed by a research community, the quality of a model becomes equal to the strength of community's professional judgement. In case the judgement lacks in some way, so does the quality of the model.

To avoid such a trap, it might be that the highly selective strategy of WoS is a more appropriate strategy for harvesting bibliometric data, as has been argued by Sandström and Sandström (2009). It may be the case that a publication indexed in WoS is, indeed, a sign of quality, since the publishing channel has passed through the WoS filter that is independent from the great diversity in quality standards that might be identified in educational research communities all over the world. On that score, bibliometric analysis based on WoS data might provide a more community-independent sense of quality of research, whereas, employing Norwegian model, this scrupulous filter is reduced to the criteria that are shared by a particular research community.

Another design feature of the Norwegian model is the relative worth of the publication type. On this aspect I do not have evidence from research, but here I believe a commonsense argument may suffice. If the number of points that a faculty would get for an article and a book are compared, then it turns out that the effort that is needed to write one Level 2 article (3 points) is equal to 60 per cent of the effort that is needed to write a Level 1 book (5 points). If take as a starting point myself – a student aspiring to an academic career – then the above mentioned reasoning may have the consequence that it makes more sense to publish Level 2 articles than books, and indeed write off the latter altogether as a waste of time. This is what in the sense of social epistemology (Fuller, 2002) are constraints for educational research in a context of $PRFS_{GU}$.

Similarly, in the Norwegian model, an implicit priority is given to publishing in English. In the list of scholarly journals that is used in this model (and in GU), there are no Level 2 journals that publish non-English articles ('The Norwegian Register for Scientific Journals, Series and Publishers: Scientific journals, series', 2016). All these journals publish only in English. This might be considered as another constraint of educational research that is conducted in a context of $PRFS_{GU}$.

The argument about language is a two-sided argument. On the one hand, it is about the expression in a particular language. On the other hand, it is about participation in a particular discussion. First, it is reasonable to say that if I am able to communicate in written academic English as well as in written academic Latvian (my mother-tongue), then I am in a more beneficial position in relation to those who are not capable of the same. My beneficial position remains the same if I prefer writing in English, which happens to be the *lingua franca* of present academia, but my ability to express myself in my mother-tongue I value as insufficient. In such a situation, I am visible to a much greater audience. Such a beneficial position cannot be identified for those who either prefer to write in a national language or mother-tongue which is not English or who do not feel comfortable with expressing themselves in written academic English. The work of such people would either be absent for databases such as WoS or as in the case of Norwegian model, be assigned to the Level 1.

But how about the other part of the argument – the participation in a particular discussion? How otherwise does the writing in English differ from, for instance, writing in Swedish? A common argument has been that in educational research is often linked to specific national phenomena, who might not be of interest for a wider (read – English-speaking) audience (Pring, 2004; Nederhof, 2006). Thus, if research is about, for example, Swedish teacher training, then it might, indeed, make more sense that I look for possibilities to have an article published in a journal such as '*Pedagogisk Forskning i Sverige'* ('Journal of Educational Research in Sweden'). However, would this mean that Swedish matters beyond Sweden would be regarded of low quality? My answer would be – no, but it may be the case that in order to pass the peer-review that is gate-keeping the road to international scholarly discussion, one may need to find a way how to contextualise local matters with international, global or simply – human matters. This may be a possible reason why prioritisation of

English as language of scholarly publications has faced critique from scholars of non-English-speaking communities.

In this study, I could identify increase in the annual average growth rate for peer-reviewed journal articles in English after the introduction of $PRFS_{GU}$, but it was not the case for articles in Swedish. For articles in Swedish, the growth was positive – every year there is increase n the number of peer-reviewed journal articles, but it remained unchanged after the introduction of $PRFS_{GU}$. If patterns for English and Swedish articles are explored in combination, then it seems that English is the prioritised language, but it does not seem to replace communication in Swedish. Hence it may be that the above raised concerns about possible knowledge effects are not relevant for GU context or, maybe rather, such an exploration of change in publishing patterns on aggregate level does not suggest a very obvious shift in publishing practices. The practice from 2005-2009 continues also in 2010-2014.

In a broader Swedish context, it was reported that English as the language of publishing accounted for a half of the publications for scholars at the three most active Swedish Research Council-grant-submitting universities in a period before the introduction of PRFS on national and university level (Hansen & Lindblad, 2010). This might suggest that it is not uncommon in Swedish educational research to publish in English and the change towards publishing in English is unrelated to PRFS.

In this discussion on consensus and language, I discussed the link between the conceptual diversity and possibilities to publish, the possibility to choose the language of publishing and also the potential problems of access. These ideas indicate that it may be problematic to link bibliometric performance with research quality, but they do not violate the very idea of a relationship between performance and quality. Hence, the first conclusion from this discussion is that, if there are enough arguments to claim that the community of educational research share consensus on what counts as publications, what is the relative worth of different publications and that writing in English is more valuable, then it is meaningful to use bibliometric indicators.

Earlier, I concluded that if the publishing patterns of educational research in GU (2005-2014) are rendered through the lens of $PRFS_{GU}$, then it seems that in this particular context the model does not seem to function well. Instead of increase in quality (the Level 2 publications), the publishing patterns seem to suggest merely increase in quantity (the Level 1 publications and those that are excluded from PRFS). The notion of quantity is the last that still needs interpretation. Given the problematic relationship between publications and the quality of educational research, what may be implications to the quality from the increase in the number of publications?

I remind, this discussion is based on assumption that the quality of educational research is a concept of social origin. Hence, for educational research to be considered as of quality, it has to be recognised as research and as research of quality. The same applies for publications. If publications of educational research are not recognised as such or, more importantly, if their worth is not recognised, then it does not matter how many publications there are. This is the conceptual link between publications and the quality of educational research that is neglected in bibliometrics and that I will explore in this final section of the discussion.

In educational research the sense of quality is closely linked to its social value for the various actors that might either benefit or use the particular knowledge claims. My proposition is that if the numbers for bibliometric performance goes up – the potential quality that this research has for those who may benefit or use this research goes down. The RED10 bibliometric analysis at the Faculty of Education in GU showed that from those 62 publications that were identified in WoS (2004-2009), about a half were uncited. Englund (2006) referred to an episode in Sweden where in a public debate on education in 1990s, educational research was simply ignored. If this is indeed the case, then there are no implications to the quality of educational research neither from the shifts in publishing culture, nor from higher scores in bibliometric performance indicators. The conceptual link between bibliometric indicators and any sense of what knowledge is and what is if for is hindered. Given such evidence, it may be the case that it does not matter how many publications there are in the educational research. They are not read and hence this research and its quality does not exist.

Let us assume that above given examples are merely exception, and it is not the case that all educational research is ignored. In that case, it is hard to imagine that more articles add anything to the social value. Consider an example. If a teacher wants to find a scholarly text that would help to understand one or another aspect of her professional work. There is no doubt that it is easier to do so among 30 and not 3000 articles. A counterargument might be that this is not a question of research, but of information search. True, and, indeed, in library and information sense this is one of the main concerns – what is the most efficient means to keep people informed. Hence, it could be argued that, the more there is knowledge, the more there is a need for mechanisms of distribution of this knowledge.

Another point would be along the lines of the argument by Winch et al. (2015) who argue that teachers often misunderstand educational research. Hence teachers should receive more education – that is what is proposed as a solution. I would consider a different solution. Earlier in this text, I made a note on a model of a university where research is linked to teaching. In such a model, teaching has been one of the means to distribute the most recent scholarly knowledge. In the present university (also in GU), research and teaching are separated activities of academic work. In these contexts, it is research that is prioritised, since additional time for research can be acquired through a competition for research grants. A conclusion from such a scenario is that that those who teach are those who have not been successful in this competition. It may be even argued that the research proposals of these people lack in quality. Here again, it may be enough with a commonsense argument from a student perspective. What quality of *education* can a student expect in a context where those who teach are those whose research (grant applications) lack in quality while those who are the 'best' are busy doing research? Does this, indeed, mean that the answer is in more education as Winch et al. (2015) suggest? What if we would consider that something has gone wrong in thinking about what knowledge is and consequently what is it for?

Is the purpose of more knowledge to merely contribute to the construction of funding distribution tables, university and country rankings? Along the lines of the normative orientation of social epistemology, a suggestion might be to take the number of uncited (Holmgren & Bertilsson Uleberg, 2011), ignored (Englund, 2006) or misunderstood (Winch et al., 2015) educational research as a starting point to a further discussion on what educational research is and what it aims to achieve.

In this discussion, I have raised various questions that I have left without an answer. Yes, I do not have answers for those questions, but the questions do all point to the strong suggestion that the competitiveness-driven argumentation underlying the performance-based research funding systems that I have explored in this study may turn out to be inverse to what is regarded as quality in educational research.

Having said that, I acknowledge attempts to contribute to the accountability of HE that can be identified in rationale of some models of evaluative bibliometrics (e.g. Sivertsen, 2010). I see the potential risk that the status of an epistemic authority that seems to still be present in universities can be abused and much time (and funding) can be wasted by producing research that is, indeed, of low quality and irrelevant. Yet, I do not see, and I hope that with this study I have been able to show that means such as PRFS can be used to govern and constrain researchers who may be contributing greatly to a better grasp of matters about and within education without offering substantively more reliable guarantees in relation to the overall conduct or quality of research.

Limitations and Further Research

As my final note I wish to comment upon a few limitations of this study that may be a source for further study. I begin with the limitations of the theoretical framework - Steve Fuller's social epistemology. The main drawback turned out to be the main characteristic of this social epistemology – the focus on the sociological sense of knowledge practices and the normative orientation. Fuller's idea is precisely that there can be no distinction between epistemic and social aspects of research quality– since within a social foundation for all knowledge, the epistemic is itself also necessarily social. On the one hand, such a conceptualisation enables a much more reflexive take on the

conceptualisation of changes in knowledge practices. It enables questioning, for example, the assumption that pursuit of knowledge is an end in itself. However, in the context of educational research such an open view on research and its goals does not allow a discussion of whether the claims about the low quality of educational research are of sociological or of epistemological kind. It may be that there is a reason to suggest that in educational research there are more instances than in, for example, sociology or psychology 'proper' of mistakes in methodological reasoning or some inconsistency in argumentation, but through the lens of social epistemology, I cannot engage in such exploration. In social epistemology these very considerations are indeed rendered irrelevant, but instead a substitute claim is moved into the foreground that, no matter what epistemic problems may be identified in educational research, if the research community allows such research to pass as legitimate, then that is all that may be said about the matter.

A few methodological limitations I have listed already earlier, but in general terms, I may suggest that an exploration of educational research in the way I have tried might have benefitted from two additional aspects of the PRFS context. It may have been beneficial to inquire about perceived change as experienced by researchers themselves and, possibly, inquire into their own sense of quality now – after several years of the introduction of PRFS. Similarly, bibliometric analysis could be conducted on individual level – thus acquiring a much finer-grained view on the change in these publishing patterns. In addition, in this study, the so-called *grey* literature was only part of background description, yet it may well have proved insightful to trace the emergent discourse on bibliometric performance *as* quality before the actual introduction of PRFS in that grey literature.

References

Aagaard, K. (2015). How incentives trickle down: Local use of a national bibliometric indicator system. *Science and Public Policy*, *42*(5), 725–737. http://doi.org/10.1093/scipol/scu087

Anderson, G., & Arsenault, N. (2005). *Fundamentals of Educational Research*. London, UK: The Falmer Press.

Anslagsfördelning och kostnadsdebitering för budgetåret 2015 samt planeringsramar för 2016-2017, Pub. L. No. E 2014/257 § Universitetsstyrelsen (2014). Retrieved from http://medarbetarportalen.gu.se/digitalAssets/1497/1497483_anslagsf--rdelning-och-kostnadsdebitering-2015-inkl-bilagor.pdf

Baez, B., & Boyles, D. (2009). *The politics of inquiry: education research and the 'culture of science'*. Albany, NY: State University of New York Press.

Ball, S. J. (2012). Performativity, Commodification and Commitment: An I-Spy Guide to the Neoliberal University. *British Journal of Educational Studies*, *60*(1), 17–28. http://doi.org/10.1080/00071005.2011.650940

Beach, D. (2011). Education science in Sweden: Promoting research for teacher education or weakening its scientific foundations? *Education Inquiry*, *2*(2). Retrieved from http://www.education-inquiry.net/index.php/edui/article/viewFile/21974/28718

Bebeau, M. J., & Monson, V. (2011). Authorship and Publication Practices in the Social Sciences: Historical Reflections on Current Practices. *Science and Engineering Ethics*, *17*(2), 365–388. http://doi.org/10.1007/s11948-011-9280-4

Biesta, G. (2014). Evidence based practice in education: Between science and democracy. In A. D. Reid, E. P. Hart, & M. A. Peters (Eds.), *A Companion to Research in Education* (pp. 391–400). Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-007-6809-3

Bornmann, L. (2011). Mimicry in science? *Scientometrics*, *86*(1), 173–177. http://doi.org/10.1007/s11192-010-0222-8

Bridges, D. (2009). Research quality assessment in education: impossible science, possible art? *British Educational Research Journal*, *35*(4), 497–517. http://doi.org/10.1080/01411920903111565

Bridges, D., Smeyers, P., & Smith, R. (Eds.). (2009). *Evidence-based education policy: what evidence? what basis? whose policy?* Malden, MA: Wiley-Blackwell.

Bueter, A. (2015). The irreducibility of value-freedom to theory assessment. *Studies in History and Philosophy of Science Part A*, *49*, 18–26. http://doi.org/10.1016/j.shpsa.2014.10.006

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, *12*(1), 39–46. http://doi.org/10.3152/147154403781776780

De Bellis, N. (2009). *Bibliometrics and citation analysis: from the Science citation index to cybermetrics*. Lanham, Md: Scarecrow Press.

Dees, W. (2008). Innovative scientometric methods for a continuous monitoring of research activities in educational science. *Proceedings of WIS*, 1–10.

Depaepe, M. (2002). A Comparative History of Educational Sciences: the comparability of the incomparable? *European Educational Research Journal*, *1*(1), 118–122.

Diem, A., & Wolter, S. C. (2013). The Use of Bibliometrics to Measure Research Performance in Education Sciences. *Research in Higher Education*, *54*(1), 86–114. http://doi.org/10.1007/s11162-012-9264-5

Elzinga, A. (2012). Features of the current science policy regime: Viewed in historical perspective. *Science and Public Policy*, *39*(4), 416–428. http://doi.org/10.1093/scipol/scs046

Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the Social Sciences and Humanities, 2000–2009. *Scientometrics*, *93*(2), 373–390. http://doi.org/10.1007/s11192-012-0680-2

Englund, T. (2006). New Trends in Swedish Educational Research. *Scandinavian Journal of Educational Research*, *50*(4), 383–396. http://doi.org/10.1080/00313830600823738

Ett lyft för forskning och innovation, Pub. L. No. Regeringens proposition 2008/09:50 (2008). Retrieved from

http://www.regeringen.se/contentassets/05cb6c62a34e4b37a114611a3ebcbd5b/ett-lyft-for-forskning-och-innovation-prop.-20080950
Fernández-Cano, A., & Bueno, A. (1999). Synthesizing scientometric paterns in Spanish educational research. *Scientometrics*, *46*(2), 349–367.
Forslag til vedtak av nye kanaler på nivå 1. (2016). Retrieved 20 April 2016, from http://dbh.nsd.uib.no/kanaler_old/nye_kanaler.do
Foss Lindblad, R., & Lindblad, S. (2016a). Higher Education and Research in a Steady State - On changing premises and practices for educational research in Sweden. In M. Elmgren, M. Folke Fichtelius, S. Hallsen, H. Roman, & W. Wermke (Eds.), *Att ta utbildningens komplexitet på allvar: En vänbok för Eva Forsberg* (Vol. No 138, pp. 209–226). Uppsala: Acta Universitatis Upsaliensis.
Foss Lindblad, R., & Lindblad, S. (2016b). On Relevance and Norms of Science in times of Restructuring: Educational Research in Sweden. In K. Trimmer (Ed.), *Political pressures on educational and social research* (pp. 66–76). New York, NY: Routledge.
Fuller, S. (2002). *Social epistemology* (2nd ed). Bloomington: Indiana University Press.
Fuller, S. (2003). Can Universities Solve the Problem of Knowledge in Society without Succumbing to the Knowledge Society? *Policy Futures in Education*, *1*(1), 106. http://doi.org/10.2304/pfie.2003.1.1.2
Fuller, S. (2007). *The knowledge book: key concepts in philosophy, science and culture*. Stocksfield: Acumen.
Fuller, S. (2010). *Science*. Durham: Acumen.
Furlong, J., & Lawn, M. (Eds.). (2011). *Disciplines of education: their role in the future of education research* (1st ed). Abingdon, Oxon ; New York, NY: Routledge.
Goldman, A. I. (1999). *Knowledge in a social world*. Oxford : New York: Clarendon Press ; Oxford University Press.
Gothenburg University Library: Allocation. (2015). Retrieved 4 December 2015, from http://www.ub.gu.se/publicera/bibliometri/resursfordelning/
Gradmann, S., Havemann, F., & Oltersdorf, J. (2014). Studies in Correlative Assessing of Intrinsic and Extrinsic Indicators of Quality. In I. Gogolin, F. Åström, & A. Hansen (Eds.), *Assessing Quality in European Educational Research* (pp. 60–84). Wiesbaden: Springer Fachmedien Wiesbaden. Retrieved from http://link.springer.com/10.1007/978-3-658-05969-9_1
Hammarfelt, B., & de Rijcke, S. (2015). Accountability in context: effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University. *Research Evaluation*, *24*(1), 63–77. http://doi.org/10.1093/reseval/rvu029
Hammarfelt, B., Nelhans, G., Eklund, P., & Åström, F. (2016). The heterogeneous landscape of bibliometric indicators: Evaluating models for allocating resources at Swedish universities. *Research Evaluation*, rvv040. http://doi.org/10.1093/reseval/rvv040
Hansen, M. (2015). Significant signs: a case study of citation practices in educational research. *International Journal of Research & Method in Education*, 1–18. http://doi.org/10.1080/1743727X.2015.1018888
Hansen, M., & Lindblad, S. (2010). *FORSKNINGSKOMMUNIKATION OCH PUBLICERINGSMÖNSTER INOM UTBILDNINGSVETENSKAP* (VETENSKAPSRÅDETS RAPPORTSERIE) (p. 95). Stockholm: Vetenskapsrådet. Retrieved from https://publikationer.vr.se/produkt/forskningskommunikation-och-publiceringsmonster-inom-utbild/
Hegel, G. W. F. (2013). *Phenomenology of spirit*. (A. V. Miller, Trans.) (Reprint.). Oxford: Oxford Univ. Press.
Henning, K. (2012). *Using the Norwegian model in a Swedish setting: consequences for social sciences and humanities*. Presented at the Nordforsk Workshop on Bibliometrics for the Social Sciences and Humanities. Retrieved from https://www.google.se/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiG8aKJ6JzMAhVLECwKHRlaBCsQFggcMAA&url=http%3A%2F%2Fwww.ub.gu.se%2Finfo%2Fbibliometri%2Fpresentationer%2Fhenning_SSH_workshop.pptx&usg=AFQjCNEIqoM9GMBqmq2F-7K-4TMRLsjc8w&sig2=NM1YC3Qj3F97K0bTYBhHjw

Henning, K. (2013). *Bedömning av publikationskanaler i de norska och danska bibliometriska modellerna – konsekvenser för humaniora och samhällsvetenskap* (p. 25). Göteborg: Göteborgs universitetsbibliotek.
Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, *41*(2), 251–261. http://doi.org/10.1016/j.respol.2011.09.007
Hicks, D., & Wouters, P. (2015). The Leiden Manifesto for research metrics. *Nature*, *520*, 429–431.
Holmgren, S., & Bertilsson Uleberg, G. (2011). *RED10 Research Evaluation*. Retrieved from https://gupea.ub.gu.se/handle/2077/24885
Inriktningsbeslut för budgetarbetet 2012. Sammanfattning, Pub. L. No. V 2011/190 (2011).
Jarneving, B., Henning, K., & Sandberg, C. (2012). *Bibliometriska analyser: Göteborgs universitet 2006-2010*. Retrieved from http://www.ub.gu.se/publicera/bibliometri/analys/Bibliometrisk-analys_granskning_r.pdf
Kant, I. (2004). *Prolegomena to any future metaphysics that will be able to come forward as science with selections from the Critique of pure reason*. (G. C. Hatfield, Ed.). Cambridge, UK; New York: Cambridge University Press. Retrieved from http://dx.doi.org/10.1017/CBO9780511808517
Kuhn, T. S. (1970). *The structure of scientific revolutions* ([2d ed., enl). Chicago: University of Chicago Press.
Lagemann, E. C. (2000). *An elusive science: the troubling history of education research*. Chicago: University of Chicago Press.
Lather, P. (1992). Critical frames in educational research: Feminist and post-structural perspectives. *Theory Into Practice*, *31*(2), 87–99. http://doi.org/10.1080/00405849209543529
Laudel, G., & Gläser, J. (2006). Tensions between evaluations and communication practices. *Journal of Higher Education Policy and Management*, *28*(3), 289–295. http://doi.org/10.1080/13600800600980130
Locke, J. (1990). *An essay concerning human understanding*. Raleigh, N.C.: Alex Catalogue. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1085946
Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar. *Journal of the American Society for Information Science and Technology*, *58*(13), 2105–2125. http://doi.org/10.1002/asi.20677
Mertens, D. M. (2010). *Research and evaluation in education and psychology: integrating diversity with quantitative, qualitative, and mixed methods* (3rd ed). Los Angeles: Sage.
Moed, H. F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? *Scientometrics*, *74*(1), 153–161. http://doi.org/10.1007/s11192-008-0108-1
Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, *66*(1), 81–100. http://doi.org/10.1007/s11192-006-0007-2
Oancea, A. (2005). Criticisms of educational research: key topics and levels of analysis. *British Educational Research Journal*, *31*(2), 157–183. http://doi.org/10.1080/0141192052000340198
Oancea, A. (2016). The Aims and Claims of Educational Research. In M. Hand & R. Davies (Eds.), *Education, ethics and experience: Essays in honour of Richard* (pp. 109–123). London ; New York: Routledge is an imprint of the Taylor & Francis Group, an Informa business.
Ossenblok, T. L. B., Engels, T. C. E., & Sivertsen, G. (2012). The representation of the social sciences and humanities in the Web of Science--a comparison of publication patterns and incentive structures in Flanders and Norway (2005-9). *Research Evaluation*, *21*(4), 280–290. http://doi.org/10.1093/reseval/rvs019
Osuna, C., Cruz-Castro, L., & Sanz-Menéndez, L. (2011). Overturning some assumptions about the effects of evaluation systems on publication performance. *Scientometrics*, *86*(3), 575–592. http://doi.org/10.1007/s11192-010-0312-7
Pajić, D. (2015). Globalization of the social sciences in Eastern Europe: genuine breakthrough or a slippery slope of the research evaluation practice? *Scientometrics*, *102*(3), 2131–2150. http://doi.org/10.1007/s11192-014-1510-5

Popper, K. R. (2002). *The logic of scientific discovery*. London; New York: Routledge. Retrieved from http://site.ebrary.com/id/5006148
Porter, T. M. (1994). Objectivity as Standardization: The Rhetoric of Impersonality in Measurement, Statistics and Cost-Benefit Analysis. In A. Megill (Ed.), *Rethinking objectivity* (pp. 197–237). Durham: Duke University Press.
Price, D. J. de S. (1971). *Little science, big science* (3. paperback print). New York: Columbia Univ. Press.
Pring, R. (2004). *Philosophy of Educational Research* (2nd ed.). London, UK: Continuum.
Pring, R. (2015). *Philosophy of educational research* (Third edition). London ; New York: Bloomsbury Academic.
Procedures for processing new submissions. (n.d.). Retrieved 20 April 2016, from https://dbh.nsd.uib.no/publiseringskanaler/OmProsedyrer?request_locale=en
Publication types in GUP. (2015). Retrieved 26 January 2016, from http://www.ub.gu.se/publicera/publikationsbas/publtyper/
Punch, K. (2005). *Introduction to social research: quantitative and qualitative approaches* (2nd ed). London ; Thousand Oaks, Calif: SAGE.
Reid, A. D., Hart, E. P., & Peters, M. A. (Eds.). (2014). *A Companion to Research in Education*. Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-007-6809-3
Rijcke, S. de, Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2015). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, rvv038. http://doi.org/10.1093/reseval/rvv038
Rosengren, K. E., & Öhngren, B. (Eds.). (1997). *An evaluation of Swedish research in education*. Uppsala: Swedish Council for Research in the Humanities and Social Sciences (HSFR).
Sandström, U., & Sandström, E. (2009). The field factor: towards a metric for academic institutions. *Research Evaluation*, *18*(3), 243–250.
Schaffner, A. (1994). The future of scientific journals: Lessons from the past. *Information Technology and Libraries*, *13*(4), 239–248.
Schneider, J. W. (2009). An Outline of the Bibliometric Indicator Used for Performance-Based Funding of Research Institutions in Norway. *European Political Science*, *8*(3), 364–378. http://doi.org/10.1057/eps.2009.19
Sismondo, S. (2010). *An introduction to science and technology studies* (2nd ed). Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.
Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI NEWSLETTER*, *6*(1), 22–28.
Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, *107*(2), 357–368. http://doi.org/10.1007/s11192-016-1845-1
Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: an empirical analysis of the potential. *Scientometrics*, *91*(2), 567–575. http://doi.org/10.1007/s11192-011-0615-3
SOCIAL SCIENCES CITATION INDEX - EDUCATION & EDUCATIONAL RESEARCH - JOURNAL LIST. (2016). Retrieved 8 May 2016, from http://science.thomsonreuters.com.ezproxy.ub.gu.se/cgi-bin/jrnlst/jlresults.cgi?PC=SS&SC=HA
SOCIAL SCIENCES CITATION INDEX - EDUCATION, SPECIAL - JOURNAL LIST. (2016). Retrieved 8 May 2016, from http://science.thomsonreuters.com.ezproxy.ub.gu.se/cgi-bin/jrnlst/jlresults.cgi?PC=SS&SC=HE
Sverige, & Justitiedepartementet. (2006). *Personal data protection: information on the Personal Data Act*. Stockholm: Ministry of Justice.
Sverige, & Regeringskansliet. (2009). *Public access to information and secrecy act: information concerning public access to information and secrecy legislation, etc*. Stockholm: Ministry of Justice, Regeringskansliet.
Swedish Higher Education Authority. (2014). *Higher education in Sweden: 2014 status report* (No. 2014:10) (p. 64). Stockholm. Retrieved from

http://english.uka.se/download/18.7ff11ece146297d1aa65b4/1407759224422/higher-education-in-Sweden-2014-status-report.pdf
Testa, J. (2016). The Thomson Reuters Journal Selection Process. Retrieved 5 May 2016, from http://wokinfo.com/essays/journal-selection-process/
The Ministry of Education and Cultural Affairs. The Act concerning the Ethical Review of Research Involving Humans, Pub. L. No. 2003:460 (2003). Retrieved from http://www.epn.se/media/1205/the_ethical_review_act.pdf
The Norwegian Register for Scientific Journals, Series and Publishers: Scientific journals, series. (2016). Retrieved 16 April 2016, from https://dbh.nsd.uib.no/publiseringskanaler/AlltidFerskListeTidsskriftSomCsv?request_locale=en
Torgerson, C. J., & Torgerson, D. J. (2001). The Need for Randomised Controlled Trials in Educational Research. *British Journal of Educational Studies*, *49*(3), 316–328. http://doi.org/10.1111/1467-8527.t01-1-00178
University of Gothenburg, & Universitetslednings kansli. Fördellning av tillkomande anslag för forskning 2009, § 10 (2008). Retrieved from http://medarbetarportalen.gu.se/digitalAssets/1417/1417612_Fordelning_av_tillkommande_anslag_2009.pdf
Valentin Kvist, A., & Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence*, *36*(5), 422–436. http://doi.org/10.1016/j.intell.2007.08.004
van Dalen, H. P., & Henkens, K. (2012). Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. *Journal of the American Society for Information Science and Technology*, *63*(7), 1282–1293. http://doi.org/10.1002/asi.22636
Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, *36*(2), 299–316. http://doi.org/10.1080/01411920902919257
Verleysen, F. T., & Engels, T. C. E. (2014). Internationalization of peer reviewed and non-peer reviewed book publications in the Social Sciences and Humanities. *Scientometrics*, *101*(2), 1431–1444. http://doi.org/10.1007/s11192-014-1267-x
Vetenskapsrådet. (2011). *Good Research Practice: The Swedish Research Council's expert group on ethics*. Stockholm: Vetenskapsrådet.
Vislie, L., Popkewitz, T., & Bjerg, J. (1997). Pedagogy and Educational Research in Sweden - History and Politics. In K. E. Rosengren & B. Öhngren (Eds.), *An evaluation of Swedish research in education* (pp. 20–41). Uppsala: Swedish Council for Research in the Humanities and Social Sciences (HSFR).
Vries, G. de. (1995). *De ontwikkeling van wetenschap: een inleiding in de wetenschapsfilosofie*. Groningen: Wolters-Noordhoff.
Wallin Peterson, M., Weibull, L., & Elofsson, I. (n.d.). Förändring för kvalitet och förnyelse. Forsknings- och utbildningsstrategier 2009–2012. University of Gothenburg. Retrieved from http://www.gu.se/digitalAssets/1039/1039508_Fus.pdf
Welch, A. (2015). Evidence-Based Policy: Epistemologically Specious, Ideologically Unsound. In *Controversies in Education* (pp. 63–79). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-08759-7_6
Whitley, R. (2007). Changing Governance of the Public Sciences. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences: the advent of research evaluation systems* (pp. 3–27). Dordrecht, the Netherlands: Springer.
Winch, C., Oancea, A., & Orchard, J. (2015). The contribution of educational research to teachers' professional learning: philosophical understandings. *Oxford Review of Education*, *41*(2), 202–216. http://doi.org/10.1080/03054985.2015.1017406

Appendices

Appendix 1 Categories for types of publications

Type in English	Type in Swedish	Description
Scientific journal article - peer reviewed	*Artikel, refereegranskad vetenskaplig*	States that the article has undergone reviews by referees.
Scientific journal article - review article	*Artikel, forskningsöversikt*	States a review article that has undergone reviews by referees.
Scientific journal article - non peer reviewed	*Artikel, övrig vetenskaplig*	States that the content of this article is of a scientific nature and directed towards the academic society but has not undergone review by referees.
Other article	*Artikel, övrig*	States that the content of this article is popular scientific and directed towards the general public, published for instance in a newspaper. Also polemical articles are included here.
Book Review	*Artikel, recension*	States book reviews in journals or newspapers.
Book	*Bok*	Monographic publication made up of one or many physical parts.
Book – peer reviewed	*Bok, refereegranskad*	Monographic publication made up of one or many physical parts that has undergone a review process by independent referees
Book - edited	Bok, med redaktör	Book with an editor. Anthology or compilation without a personal author.
Book, edited – peer reviewed	*Bok med redaktör, refereegranskad*	Book with an editor that has undergone a review process by independent referees. Anthology or compilation without a personal author.
Book Chapter	*Kapitel*	Independent part of a monographic publication, often a compilation.
Book Chapter - peer reviewed	*Kapitel, refereegranskat*	Independent part of a monographic publication (often a compilation) that has undergone a review process.
Doctoral thesis	*Doktorsavhandling*	Passed thesis for doctoral degree (Swedish).
Licentiate thesis	*Licentiatavhandling*	Passed thesis for licentiate degree (Swedish).
Conference Contribution	*Konferensbidrag, övrigt*	Refers to contribution such as oral presentation, workshop, or non-peer reviewed paper
Conference Poster	*Konferensbidrag, poster*	Refers to contribution to a poster session
Conference Abstract - peer reviewed	*Konferensbidrag - refereegranskat abstract*	Refers to a peer reviewed summation of a completed paper
Conference Paper - peer reviewed	Konferensbidrag, refereegranskat	Refers to paper reviewed by independent referees and published in proceedings with ISSN or ISBN
Report	*Rapport*	Refers to a publication which is a part of a report series.
Artistic research and development project	*Konstnärligt forsknings- och utvecklingsarbete*	Documentation from an artistic research and/or development project. Applicable up to and including 2010.
Textbook.	*Lärobok*	Book intended primarily for teaching purposes. The content is aimed at student at colleges and universities
Other	*Övrigt*	To be used only for publications where no other type is applicable. Also translations are included here.

('Publication types in GUP', 2015)

Appendix 2 Excluded and transformed categories for types of publications

Original	N	Result
Book chapter, peer-reviewed	126	Combined with "Book chapter"
Other	43	Excluded
Book - edited, peer-reviewed	8	Combined with "Book - edited"
Book, peer-reviewed	4	Combined with "Book"
Artistic research and development project	1	Excluded

Appendix 3 Variables

Label	Name	Description	Level	Values
ID	pubid	Identification number of a publication	categorical	numerical
Year	y	Publication year	continuous	1-2005; ... ; 10-2014
Type	t	Publication type	categorical	1 Scientific journal article - review article 2 Other article 3 Scientific journal article - non peer reviewed 4 Book Review 5 Scientific journal article - peer reviewed 6 Book 7 Book – edited 8 Doctoral thesis 9 Book chapter 10 Conference abstract - peer reviewed 11 Conference contribution 12 Conference poster 13 Conference paper - peer reviewed 14 Textbook 15 Licentiate thesis 16 Report
Language	lang	Publication language	categorical	0-Swedish 1-English 2-Other
Level	nor	Level in the bibliometric model	categorical	0-not included 1-Level 1 2-Level 2

Appendix 4 Abbreviations used in the statistical tables

P	Total number of publications, full counts
$P_{2005\text{-}2009}$	Number of publications before the introduction of PRFS, 2005-2009, full counts
$P_{2010\text{-}2014}$	Number of publications after the introduction of PRFS, 2010-2014, full counts
PR	Percetage growth rate in the number of publications after the introduction of PRFS in relation to the number before
$C_{2005\text{-}2009}$	Average annual growth rate before the introduction of PRFS (2005-2009)
$C_{2010\text{-}2014}$	Average annual growth rate after the introduction of PRFS (2010-2014)
$P_{2005} \ldots P_{2014}$	number of publications per year, full counts

Appendix 5 Descriptive statistics for all publications: type, year, PRFS

Table continues (page 1 of 2)

Type	PRFS	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}
Scientific journal article - peer reviewed	*All*	753	248	505	104%	58%	65%	40	52	42	54	60	68	90	88	122	137
	All included	707	230	477	107%	57%	66%	38	47	37	50	58	63	90	80	113	131
	Level 1	544	162	382	136%	57%	71%	27	29	29	39	38	45	68	66	98	105
	Level 2	163	68	95	40%	58%	52%	11	18	8	11	20	18	22	14	15	26
	Excluded	46	18	28	56%	73%	54%	2	5	5	4	2	5	0	8	9	6
Scientific journal article - non peer reviewed	*All*	156	70	86	23%	67%	41%	9	4	14	19	24	22	20	14	17	13
	All included	1	1	0	-	-	-	0	1	0	0	0	0	0	0	0	0
	Level 1	1	1	0	-	-	-	0	1	0	0	0	0	0	0	0	0
	Excluded	155	69	86	25%	66%	41%	9	3	14	19	24	22	20	14	17	13
Scientific journal article - review article	*All*	34	7	27	286%	63%	52%	1	0	0	2	4	5	6	8	8	0
	All included	21	1	20	1900	-	61%	0	0	0	0	1	3	5	8	4	0
	Level 1	20	1	19	1800	-	59%	0	0	0	0	1	3	4	8	4	0
	Level 2	1	0	1	-	-	-	0	0	0	0	0	0	1	0	0	0
	Excluded	13	6	7	17%	57%	37%	1	0	0	2	3	2	1	0	4	0
Book Review	*All**	113	35	78	123%	105%	41%	2	2	13	3	15	20	21	12	17	8
Other article	*All**	301	131	170	30%	90%	61%	10	32	32	22	35	25	33	23	50	39
Book	*All*	130	56	74	32%	75%	49%	6	14	9	18	9	15	16	14	16	13
	All included	88	41	47	15%	62%	61%	6	10	5	14	6	7	9	12	8	11
	Level 1	82	40	42	5%	61%	63%	6	10	4	14	6	6	8	11	8	9
	Level 2	6	1	5	400%	-	50%	0	0	1	0	0	1	1	1	0	2
	Excluded	42	15	27	80%	-	36%	0	4	4	4	3	8	7	2	8	2
Book-edited	*All**	109	40	69	73%	78%	71%	4	11	5	5	15	8	22	13	14	12
Book chapter	*All*	852	377	475	26%	60%	72%	57	75	70	68	107	54	112	86	106	117
	All included	565	257	308	20%	68%	71%	32	44	44	45	92	36	89	66	50	67
	Level 1	484	223	261	17%	70%	72%	27	39	35	42	80	30	84	56	42	49
	Level 2	81	34	47	38%	61%	67%	5	5	9	3	12	6	5	10	8	18
	Excluded	287	120	167	39%	48%	75%	25	31	26	23	15	18	23	20	56	50
Type	PRFS	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}

*All are excluded from PRFS

Table continued (page 2 of 2)

Type	PRFS	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}
Textbook	*All*	21	8	13	63%	68%	34%	1	2	2	1	2	4	2	0	5	2
	All included	2	2	0	-100%	-	-	0	2	0	0	0	0	0	0	0	0
	Level 1	2	2	0	-100%	-	-	0	2	0	0	0	0	0		0	0
	Excluded	19	6	13	117%	57%	34%	1	0	2	1	2	4	2		5	2
Reports	*All*	209	162	47	-71%	42%	35%	40	45	29	33	15	14	6	13	4	10
Conference abstract - peer reviewed *	*All*	185	10	175	1650%	35%	124%	3	0	3	3	1	7	6	12	55	95
Conference contribution*	*All*	389	185	204	10%	54%	75%	33	47	35	44	26	22	27	46	67	42
Conference poster*	*All*	55	18	37	106%	57%	47%	3	3	4	3	5	8	9	10	7	3
Conference paper - peer reviewed	*All*	645	333	312	-6%	57%	46%	55	49	52	107	70	69	90	88	48	17
	All included	49	11	38	245%	82%	109%	1	2	3	1	4	2	7	5	12	12
	Level 1	48	10	38	280%	-	109%	0	2	3	1	4	2	7	5	12	12
	Level 2	1	1	0	-100%	-	-	1	0	0	0	0	0	0	0	0	0
	Excluded	596	322	274	-15%	56%	42%	54	47	49	106	66	67	83	83	36	5
Doctoral thesis	*All*	119	58	61	5%	55%	47%	10	11	9	13	15	13	15	8	16	9
	All included	62	33	29	-12%	-	76%	0	0	8	10	15	3	12	7	3	4
	Level 1	62	33	29	-12%	-	76%	0	0	8	10	15	3	12	7	3	4
	Excluded	57	25	32	28%	26%	34%	10	11	1	3	0	10	3	1	13	5
Licentiate thesis*	*All*	19	0	19	-	-	48%	0	0	0	0	0	4	2	2	1	10
Type	*PRFS*	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}

*All are excluded from PRFS

Appendix 6 Descriptive statistics for scientific journal articles – peer reviewed: language, year, PRFS

Language	PRFS	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}
All languages	*All*	753	248	505	104%	58%	65%	40	52	42	54	60	68	90	88	122	137
	All included	707	230	477	107%	57%	66%	38	47	37	50	58	63	90	80	113	131
	Level 1	544	162	382	136%	57%	71%	27	29	29	39	38	45	68	66	98	105
	Level 2	163	68	95	40%	58%	52%	11	18	8	11	20	18	22	14	15	26
	Excluded	46	18	28	56%	73%	54%	2	5	5	4	2	5	0	8	9	6
In English	*All*	655	209	446	113%	61%	68%	31	47	37	46	48	56	83	73	105	129
	All included	621	197	424	115%	59%	70%	31	43	33	43	47	51	83	69	97	124
	Level 1	460	130	330	154%	58%	78%	21	25	25	32	27	33	62	55	82	98
	Level 2	161	67	94	40%	61%	51%	10	18	8	11	20	18	21	14	15	26
	Excluded	34	12	22	83%		45%	0	4	4	3	1	5	0	4	8	5
In Swedish	*All*	86	33	53	61%	47%	48%	7	5	4	7	10	11	7	12	15	8
	All included	80	31	49	58%	45%	45%	7	4	3	7	10	11	7	10	14	7
	Level 1	78	30	48	60%	50%	45%	6	4	3	7	10	11	6	10	14	7
	Level 2	2	1	1	0%	-	-	1	0	0	0	0	0	1	0	0	0
	Excluded	6	2	4	100%	-	-	0	1	1	0	0	0	0	2	1	1
In other languages	*ALL*	12	6	6	0%	-	-	2	0	1	1	2	1	0	3	2	0
	Included	6	2	4	100%	32%	57%	0	0	1	0	1	1	0	1	2	0
	Level 1	6	2	4	100%	-	41%	0	0	1	0	1	1	0	1	2	0
	Excluded	6	4	2	-50%	-	41%	2	0	0	1	1	0	0	2	0	0
Language	PRFS	P	$P_{2005-2009}$	$P_{2010-2014}$	PR	$C_{2005-2009}$	$C_{2010-2014}$	P_{2005}	P_{2006}	P_{2007}	P_{2008}	P_{2009}	P_{2010}	P_{2011}	P_{2012}	P_{2013}	P_{2014}