

# **D 2.2.1.4**

# **Report on the Use and Non-Use of Indicators in the SSH**

Three Focal Points for the Case of the  
Social Sciences and Humanities

## Table of Contents

Table of Contents	2
Document Information	3
1. Introduction	4
<b>1.1. What are indicators?</b>	<b>5</b>
2. Cases of indicator use	7
<b>2.1. Italy – The Evaluation of Research Quality (VQR)</b>	<b>7</b>
<b>2.2. Denmark Bibliometric Research Indicator (BRI)</b>	<b>8</b>
<b>2.3. United Kingdom – Research Excellence Framework (REF)</b>	<b>8</b>
3. Presentation of literature	9
<b>3.1. Contextualising indicators</b>	<b>9</b>
<b>3.2. Focus points</b>	<b>10</b>
3.2.1. Critical, novel, creative research, and indicators	11
3.2.2. The balance between education and research	13
3.2.3. The special case of the social sciences and humanities	15
4. Discussion: Consequence for reform-type universities	17
5. Conclusion	18
6. References	19

The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

## Document Information

<b>Grant Agreement</b>	101004053
<b>Project Title</b>	European Reform University Alliance
<b>Project Acronym</b>	ERUA
<b>Project Start Date</b>	01 November 2022
<b>Related Work Package</b>	WP 2 – Re-Imagining Higher Education and Research
<b>Related Tasks</b>	Deliverable 2.2.1.3
<b>Lead Beneficiary</b>	Roskilde University
<b>Submission Date</b>	01 November 2022
<b>Dissemination Level</b>	Public
<b>Authors</b>	P8: .....; RUC: .....; UKon: .....; UAegean: .....; NBU: .....

## 1. Introduction

Whereas D.2.2.1.2 focused on the rising use and institutionalization of societal impact measurements – societal, economic, social, and so forth – this report addresses scientific indicator measures used in the social sciences and humanities. In other words, where societal impact measures broadly address measures and evaluations external to research and the academic world itself, this chapter zooms in on the internal research measures and evaluations – i.e., those situated *within* the academic and scholarly world and function on the premises of a scholarly logic.

Thus, in this report, we present an overview of the literature on the issue of the use and non-use of indicators in the social sciences and humanities. We focus on three key issues of the use of indicators that are relevant to ERUA: 1) how indicator use affects the conditions for critical, novel, creative research, 2) how it affects the balance between education and research at universities and finally, 3) how it impacts the social sciences and humanities in particular. The report concludes with a summary of selected recommendations extracted from the relevant literature. In this report, we review the existing literature and evaluation systems through the lens of the overarching ideas and values of ERUA and specifically WP2 (Reimagining Higher Education and Research). Thus, our focus is oriented towards the academic consequences of evaluation systems that relate to creativity, experimentation, and academic freedom.

Reviews of the existing literature on the broad contexts of effects of impact and indicator use has been carried out previously (De Rijcke et al., 2016; Grant et al., 2010). However, firstly, years have passed since these have been carried out, and we argue that in a field infused with policy changes and changing empirical realities, an updated examination is needed. Secondly, we identify a need for further (critical) examination of the field in the context of the scientific values and research goals put forward by ERUA. This also means that we will view the existing literature through the lens of a Social Sciences and Humanities (SSH) perspective.

Thus, the goals of the report are the following:

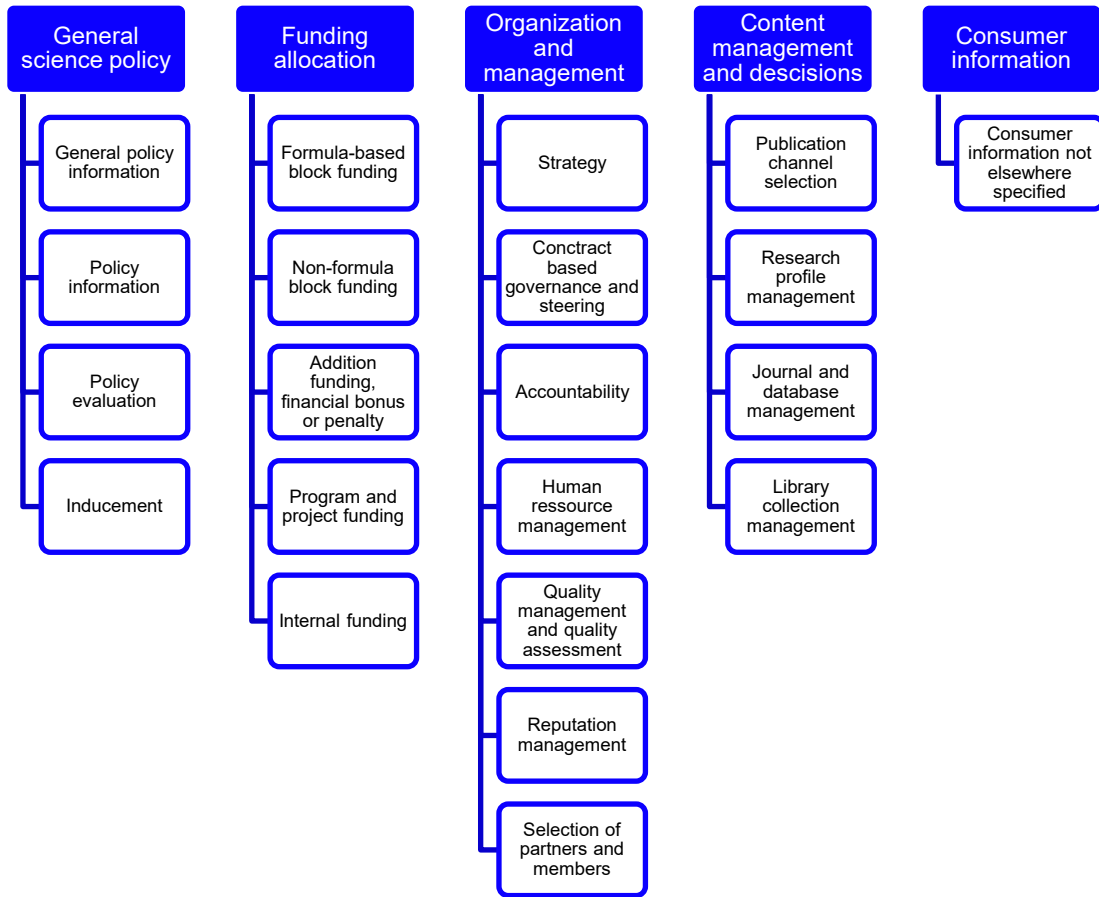
1. Present an updated overview of cases of indicator use and non-use with a special focus on Social Sciences and Humanities

2. Present literature based on ERUA values and need for further research in relation to the issue of indicator use and non-use specifically in the Social Sciences and Humanities
3. Deduce recommendations for the use and non-use of indicators in relation to ERUA's mission and values

### 1.1. What are indicators?

Before proceeding to the presentation of the literature, some theoretical clarity is needed. In this report, we use a broad conceptualization of indicator use. Thus, what we refer to comprises most aspects of bibliometrics, i.e., the quantitative analysis of publications and citations, but in this report, we do not include measures of 'external' scientific value or quality, i.e., the perceived impact that a publication or project has outside of academia itself, the societal impact value. In addition to drawing on the indicators themselves, we also extend our focus to perhaps the most substantial and direct (financial and organizational) consequences, the case of performance-based research funding systems.

Kosten (2016) constructed a general overview of research indicator use, which presents 21 categories of indicator use which can broadly be grouped into five main categories. His classification is presented in Figure 1. The tables highlight that the use of indicators is broad and covers many areas of scientific practice and administration – and although funding allocation is perhaps the most discussed issue in the literature, indicator use informs many other aspects of scientific life.



## 2. Cases of indicator use

In order to contextualise the subject of this report, in this section, we present contemporary cases of use and non-use of scientific performance indicators. As shown, indicators can be used in an array of ways, however, here we focus specifically on the case of so-called ‘Performance-based Research Funding Systems’ (PRFS). These systems, although their specific characteristics vary drastically, all share the same underlying principle of allocating national research funds based on some form of evaluation of research at different unit levels (Hicks, 2012, p. 252). There are a multitude of rationales pushing the movement towards the implementation and use of these systems, including increasing productivity without the addition of funds as well as the desire for universities to be based on market incentives and public service to a higher degree (Hicks, 2012, p. 253). In addition, the systems are widely used: Jonkers & Zacharewicz (2016, p. 6) examined the use of PRFS’s in the EU and found that most member states had implemented some variant of these types of systems.

We present three cases: Italy’s ‘VQR’, Denmark’s ‘BRI’, and UK’s ‘REF’. These models constitute different approaches to the measurement of performance and the allocation of funds on the basis of measurement. VQR and REF are so-called ‘panel based models’, where reviewers draw on bibliometric measures to review various aspects of researchers’ performance, and BRI is a ‘publication based model’ where the unit of assessment is publications. In addition, some countries (Poland, Slovakia, Sweden) also use ‘citation based models’ where the frequency of citation is used as the unit of assessment (Mouritzen et al., 2018, pp. 22–23).

### 2.1. Italy – The Evaluation of Research Quality (VQR)

In 2011, Italy launched its research assessment system, the VQR (Valutazione della Qualità della Ricerca) which combines peer review and bibliometric methods (Ancaiani et al., 2015). Peer review is employed in the SSH while bibliometric methods are used in the STEM disciplines (Bonaccorsi, 2020). The evaluation is carried out every five years (Italian National Agency for the Evaluation of Universities and Research Institutes, n.d.) by the Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANUVUR). In the 2004-10 research evaluation exercise, 14 Groups of Evaluation Experts (GEV), one for each research area, evaluated the research performance of Italian universities and research bodies. Italian researchers submitted their best publications to the universities and research

bodies which then made the final selection of which work should serve as the basis for assessment. The GEV rated the submissions against the criteria of relevance, originality and innovation, and internationalization on a scale from excellent to limited (Ancaiani et al., 2015).

## 2.2. Denmark Bibliometric Research Indicator (BRI)

The BRI was a Danish model for allocating research funds for universities based on the performance of their total publishing performance initiated in 2006 and implemented in 2008 (Mouritzen et al., 2018, p. 34). The system was retired at the end of 2021 on the basis of a broad political agreement which altered the funding structure significantly. In this model, 67 discipline-specific expert subject groups assess and classify journals into three levels corresponding to number of points per publication in these channels: 1 (normal), 2 (high), and 3 (excellent). In addition to publications in peer reviewed scientific journal, scientific monographs are also rewarded with points according to the publisher (textbooks for students and practitioners are not included). On the basis of the universities' total publications and corresponding points, they are rewarded a share of the funds anchored in the system. In total, 25 % of the 'basic funds' were allocated on the basis of the performance measured in BFI points (Ministry of Higher Education and Science, 2019, 2020; Mouritzen et al., 2018, pp. 23–33). In the political agreement to end the current use of the BFI system, the heavy administrative and financial burden was emphasized: 430 researchers were involved in the continuous work to maintain the BFI lists, and it was argued that the end of BFI constituted a 'clear administrative simplification for the universities and will free time for the involved researchers' (Political agreement, p. 3). However, the use of performance to allocate basic funds in the Danish university sector continues: The BFI has been 'frozen in place' so as to allocate funds based on the latest list of publication channels, although the share will be lower (Socialdemokratiet et al., 2021).

## 2.3. United Kingdom – Research Excellence Framework (REF)

The first PRFS to be introduced was the Research Assessment Exercise (RAE) in the UK, which was initiated in 1986 and was replaced by the current UK system, the Research Excellence Framework (REF) in 2014 (Hicks, 2012, p. 252). Contrary to many other systems, the REF is not merely based on metric exercises, but rather on expert peer reviews carried out in 34 subject-based assessment units, headed by four main panels. These reviews take



place every sixth or seventh year. The experts in these panels and units comprise senior academics, international members and so called ‘research users’ from both private and public sectors<sup>1</sup>. Based on evaluation in REF, approximately £ 2 billion (€ 2,3 billion) of public funds is allocated to universities. Submissions to the REF are assessed by the quality of outputs (60 % of the total score), their impact ‘beyond academia’ (25 % of the total score) as well as the environment that supports research (15 % of the total score) (Higher Funding Council of England England, n.d.).

### 3. Presentation of literature

In this section, we present a brief overview of relevant literature on the issue of use and non-use of indicator in the social sciences and humanities. We view the literature through the lens of the overarching ideas and values constituting the basis of ERUA, namely the furthering of critical and novel research, interdisciplinarity, academic experimentation and freedom and so forth.

#### 3.1. Contextualising indicators

The rising impact of indicators in the scientific community is part of a larger societal shift towards a greater focus on evaluations and effect measures within the last decades: the trend has even been referred to as a ‘culture of evaluation’ (Dahler-Larsen, 2011) and has been linked especially to the theoretical administrative paradigm ‘new public management’, which emphasizes the core values of ‘incentivization’ and competition in funding allocation (Dunleavy et al., 2006). Thus, the shift to funding structures based partly on performance measured through various indicators for productivity and quality echoes a larger administrative shift, focusing on increased ‘*efficiency, productivity and applicability*’ (Auranen & Nieminen, 2010; Wang et al., 2018). And though the European, Japanese and North American public scientific systems have all been traditionally characterized by autonomy and limited emphasis on evaluations, they have also been severely affected by the ‘evaluation wave’, specifically when it comes to funding structures (Wang et al., 2018, p. 3).

---

<sup>1</sup> See <https://www.ref.ac.uk/about-the-ref/what-is-the-ref/>

There is generally agreement that various performance evaluation systems – most often in relation to the allocation of research funds – have gained importance in recent decades (Aagaard et al., 2015; Hicks, 2012; Jonkers & Zacharewicz, 2016). In addition, the importance of being aware of indicator systems relates directly to the mission and values of ERUA: At the heart of the project lies an underlying set of values that points out universities as institutions of experimentation; as “*creative spaces*”, where disruption is likely to and *should* occur. Likewise, the ERUA mission underlines interdisciplinary development as a key goal. However, as we will show, the potential effect of evaluation systems on exactly these aspects of science; on experimentation, on creativity, is up for discussion within the scholarly literature.

### 3.2. Focus points

To limit the presentation of literature, we take our point of departure in three distinct focal points relating to the founding values of ERUA. In this section, we present each of these focal points. A cross-cutting theme of these is that they address the underlying question of whether the use of indicators in various ways increases strategic behavior of researchers, i.e., if they alter scientific production in ways that are related to a desire to perform in a specific way based on them.

Why are the introduction of indicators expected to influence the behaviour of researchers? The introduction of strategic responses to performance measures have been theorized in many ways, and one approach that offers insight into these types of systems is offered by the notion of ‘effort substitution’, i.e., the production of improvement in performance in the indicator that is being measured – i.e., in what *can* and *is* being measured on the basis of various incentives such as allocation of research funds – at the expense of other measures or tasks (Kelman & Friedman, 2009, p. 922). In addition, Bonaccorsi (2020, p. 10) illustrates the behavioral impact that various types of indicator use in research might have by pointing out that indicators live a life of their own after their initial implementation: They do not merely measure and quantify research, they ‘*create incentives and disincentives, recommend some behaviors and discourages others*’ (Bonaccorsi, 2020, p. 10). The specific nature of these patterns of changes in scientific behavior and strategy is the subject of this document. In essence, the underlying principle that the types of systems that are the subject of this report might change the basic incentives structuring the everyday working life and priorities of university researchers.

It is worth noting that the establishment of causal evidence of the nature and extent of consequences of performance-based research funding systems faces several methodological challenges, which will be evident throughout the presentation. It is particularly hard to attribute causal effects, especially seen in the light of the typical existence of multiple systems set in place to enhance productivity, quality, strategizing and so forth at university, regional, national and supra-national levels (OECD & Butler, 2010, p. 128).

### 3.2.1. Critical, novel, creative research, and indicators

At the heart of the ERUA projects lies the notion that research must be able to be critical, that universities should be creative hubs of novel and experimental approaches to science. However, one might fear that an increased use of performance indicators might lead to a strategic behavior of scientists in terms of choosing more ‘safe’ research subjects and approaches that are more likely to result in successful publications.

Generally, while there is not yet final agreement on the empirical reality of the causal relation between funding systems and general productivity (Geuna & Martin, 2003), the question has been explored rather extensively in the field (Cattaneo et al., 2016). However, the more specific question of how these structures affect the level of creativity and experimentation is under-examined. Additionally, the studies that do exist seldom draw on empirical evidence showing the causal relation between strategies relation to critique, novelty, creativity, and experimentation and increasing use of indicators, since it is extremely hard to measure. Thus, many studies use a theoretical or ideological approach to critique the notion of indicator use, and do not include empirical data as such (Adler & Harzing, 2009; Hammarfelt & de Rijcke, 2015; Lawrence, 2008).

In a Danish context, Mouritzen et al. (2018, pp. 204–205) point to the emergence of public debate on the BRI and the expectations that it might have negative consequences in terms of prioritising research that is ‘sure to be published’ at the cost of innovative and interdisciplinary experimentation, and some feared that it might encourage ‘assembly line research’ rather than ‘jumping into the deep end and think in new ways in relation to a problem’. This critique was particularly prevalent among scholars within the ‘dry sciences’ (SSH), and was most pronounced within faculties of humanities. However, they do not find large empirical changes in the research that is actually carried out and being published, and researchers point to other forms of incentives, such as status in academia, as important for their publishing strategy (Mouritzen et al., 2018, pp. 230–231). In a Spanish context,

Cañibano et al. (2018) show that researchers feel that their time available for development of ideas prior to publication has decreased at the expense of rigour and possibility to take part in long-term projects. Accordingly, they argue that the evaluation system in place hinders the basic ideas of creativity and originality.

In terms of whether indicator use, specifically in the context of performance based allocation of funds, impacts the conditions for creativity and experimentation, there are some studies that hypothesise that it might have a negative impact.

Wang et al. (2017) point out that ‘novel papers’ are less likely to end up in the top 1 %, which makes them more risky than other types of studies. Thus, bibliometric indicators might be biased against novel research, and the results of Wang et al. also lead to a possibility to link novelty and indicators to some extent. Likewise, drawing on principal-agent theory (Liefner, 2003) finds empirical evidence that the introduction of performance-based funding implies that researchers will be less likely to involve themselves in projects that have high chance of failure. Geuna & Martin (2003, p. 297) point to the ‘homogenization’ of research and universities as a potential drawback of funding based on indicator performance: They argue that these systems might discourage risky and innovative research, resulting in a lower level of research diversity and experimentation, and may result in a reinforcement of research elites, affirming the status quo, because it rewards *prior* performance and not the future or current potential of a research project or research institution or group. In line with this, based on a study of history departments and the British REF-system, Hamann (2016) argues that research performance assessments constitutes a form of ‘visible hand’, promoting disciplinary standardization and resource allocation based on a specific, standardized notion of ‘excellence’ rather than research performance. He argues that the REF has involved a ‘necessary reduction of complexities and uncertainties’ (Hamann, 2016, pp. 775–777), and that in the UK context, prestigious and established departments and universities monopolize reputation. Likewise, he argues that researchers might strategically adapt their publications to fit the narrow definition of excellent research – that which is published in the highest ranking journals (Hamann, 2016, pp. 775–777).

In terms of the conditions for interdisciplinarity, according to a literature review conducted by De Rijcke et al. (2016, pp. 163–164), concerns have been raised that assessment systems such as REF have a negative influence on the conditions for interdisciplinary research. In addition, some raise concerns over the issue that various bibliometric indicators, which are used in some funding systems, might have trouble taking into account interdisciplinary

research, because it often uses various forms of standardization to account and correct for subject or discipline specific practices (Ávila-Robinson et al., 2021; Hamann, 2016).

Thus, although some studies theoretically or empirically hypothesise or illustrate that there is a possibility that the use of indicators, specifically when used to allocate research funds, has a negative impact on the type of research that ERUA aims to further and provide fruitful conditions for, that which is creative, innovative, and experimental, the issue is insufficiently examined. We call for further research into the specific nature of this impact and whether it does exist. Such research might draw on administrative analysis of research production, but assessing the self-perceived research strategies amongst scholars is equally important: What role do indicators, for example in relation to funding allocation, play in their everyday life as scientists?

### 3.2.2. The balance between education and research

A question relevant to the issue of evaluation measures in science in the context of ERUA is whether the pervasive culture of scientific evaluation – and more specifically the trend that university-level or individual funding is linked to performance on these measures – shifts the balance between research and teaching towards a lesser focus on the latter. The hypothesis that this question rests upon is that because researchers are under pressure to perform on the basis of quantitative research indicators – to publish more articles, to publish in the best – and international – journals, to get more citations or to receive more and bigger grants, they are prone to neglect teaching as a task; simply because this does not play into how they are evaluated as scholars and thus has no visible or obvious effect on their scientific career prospects.

Briefly put, rewards for teaching might be smaller than those for research. In the ERUA-project, we put great focus on learning, teaching, and pedagogical innovation. But innovation and creativity in learning and teaching requires careful thought and resources from researchers and scholars. This is the lens through which we view the question of the balance between teaching and research. In this section we examine the empirical and theoretical evidence and perspectives on this dilemma.

First, like many of the issues interesting to us, the topic is insufficiently examined. Although the balance between teaching and research – the so-called ‘teaching-research nexus’ – has been the subject of a number of studies (De Witte et al., 2013; Hattie & Marsh, 1996;

Neumann, 1992), the extent of studying the connection to funding mechanisms and indicators has been very limited. However, research indicator systems typically do not involve the measurement of educational quality, although it might be argued that this is one of the most important roles of modern universities, seeing as the educational background of university graduates will have a large impact in society through employment. To contextualise the issue, it is worth mentioning that Coate et al. (2001, p. 172) point out that teaching and research are seen as distinct activities in terms of how they are managed; and often, because of limited resources, the two might compete for attention and dedication.

Geuna & Martin (2003, p. 297) point to the separation of teaching and research as a drawback of performance-based research funding, i.e. funding based upon how well institutions perform based on the indicators that are the subject of this review. According to them, the consequence of this separation is that teaching will be prioritized lower than research, because the reward for research is greater than that of teaching. However, they offer limited empirical evidence on these thoughts, and they are theoretical rather than data based.

Additionally, Mägi & Beerkens (2016, p. 242) argue that the marketisation and globalization and more specifically the idea of the university as a 'single good' expressed in terms of research excellence, has pulled research and teaching apart. Likewise, they point out the increasing tendency that individual researchers and universities are prioritizing teaching to a lesser extent, for example by using research grants to 'buy out' of teaching activities or by an increasing share of teaching being dealt with by adjunct/non-tenured staff that do not take part in research activities, which conflicts with the idea of a holistic combination of teaching and research.

In the Danish context, the recently retired BRI system has been the subject of a thorough investigation covering some of the points that are specifically interesting to ERUA and the present report: Namely, Mouritzen et al. (2018) assessed the effect of the system on dissemination and teaching and find that there is actually a negative correlation between time spent on research and the implementation of the BRI, and likewise, a positive correlation with time spent on teaching, and thus did not find that teaching was affected negatively.

We also call for further research into the relation between, on one hand, teaching and research, and on the other, indicator use and systems: Once again, studying the perceptions of this correlation amongst academics in systems that are influenced by indicators might be

particularly fruitful: How do researchers think about this balance, and how do they go about planning and devoting time to teaching in their everyday work? And how do they assess the meaning of each in terms of their future career?

### 3.2.3. The special case of the social sciences and humanities

Finally, we will discuss whether there is any literature suggesting that there are special characteristics of indicator use and non-use in the social sciences and humanities that might be interesting for ERUA.

Generally, while the introduction of indicator use and citations-based quantitative assessment is not a new concept in STEM, quantitative measures have been met with strong opposition in the SSH, specifically in the humanities (Ochsner et al., 2016).

Perhaps the most obvious point to discuss when examining the use and non-use of research indicators in the social sciences and humanities specifically is the issue of publications language. In general, in comparison to STEM, SSH research is much more regionally or nationally anchored, and in non-Anglosaxon countries, more likely to be published in national languages, which will often not result in a high score in most indicators, since they typically favour English-language channels (Mathies et al., 2020, p. 22; Melchiorson, 2019, p. 367; Nederhof, 2006, pp. 83–84). Why might it be a problem that increased focus on indicators pushes the share of publications being published in English-language channels? (Sivertsen, 2016) 358 argues that SSH will lose their *raison d'être* if publication is exclusively international, because they must stay connected with the surrounding society and cultural context.

In the Danish performance-based funding mechanisms, the BRI, virtually no publication channels in the national language are classified as high level, and the indicator has been shown to decrease the number of publications in Danish indexed journals. Likewise, in the Finnish university sector, a similar pattern was found, and the increase in English-language publications was primarily driven by the social sciences and also to some extent the humanities (Mathies et al., 2020, p. 32), and generally SSH researchers were most affected by the implementation of performance based funding. In Norway, a moderate shift away from national language publishing was identified especially in the SSH (Aagaard et al., 2014, p. 7).

According to Ochsner et al., (2017, pp. 2–3) there are two main issues related to the application of bibliometric methods to SSH disciplines: *coverage* and *methodological* issues, which is generally echoed in the literature. Coverage issues arise because publication patterns in SSH are often different from those in STEM: In addition to language differences, researchers are more likely to cite book and monograph chapters, and most bibliometric methods are mainly based on journal articles (Nederhof, 2006, p. 84).

Finally Ochsner et al. (2017, pp. 2–3) point out that SSH scholars not only disseminate their research to other researchers but also to the broader society, which is not accounted for in bibliometric analysis (Nederhof, 2006, pp. 88–89). In terms of methodological issues, they point to the transdisciplinarity of journals as an issue when attempting to conduct ‘field normalization’, i.e., when attempting to correct bibliometric data for field- or discipline specific publication and citation patterns (Hicks et al., 2004). Nederhof (2006, p. 92) thus argues that bibliometric monitoring can in general be used for SSH with the same methods used in STEM, but that there are several extensions, such as the need to include a broader range of publications and indicators, for example publications aimed at non-academic audiences and monographs.



## 4. Discussion: Consequences for reform-type universities

In relation to 2.2.1.2 which underlines that external indicators are very hard to measure, the consequence might be that universities and scholars take ‘the easy way out’ and focus on internal impact rather than external, which is crucial to reform-type universities. We have identified some literature that either presents empirical evidence or discusses theoretical implications of the use of indicators, especially in relation to the allocation of funds for research. Although much more empirical inquiry as well as theoretical discussions of the issue are needed, we at least call for increased awareness of the implications that increasing use of indicators might have for the type of research and the institutional models and principles that are at the heart of the idea of a reform university.

As reform universities, we should be critical of the structures governing the conditions of the research and institutional characteristics that we wish to further. We have shown that there are implications that the use of indicators might lead to homogenization of research, and that it might come at the cost of prioritisation of teaching, which also implied that this might happen at the cost of interdisciplinarity. However, in order to not draw incorrect conclusions, we should be aware whether the negative implications for the type of research and the institutional layout that reform universities try to promote is inherent to the use of performance indicators or if it is rather a problem relating to the current state of indicator methodology. Thus, we might imagine an indicator system that had a completely different focus, maybe even designed to further the values of interdisciplinarity, pedagogical innovation as well as cross-collaboration.

## 5. Conclusion

Indicator use has been institutionalised in various forms of university life: In the EU, many member states have implemented performance systems based on indicators in some form. The use of indicators takes many forms, although the underlying principle of accountability and marketization is re-occurring. In terms of the focal points that we emphasized, there is some theoretical debate and empirical evidence that point to the negative consequences of an extensive use of indicators when examining the conditions for the type of research and university structure that ERUA seeks to further: Namely one that affords the possibility for novelty, creativity, experimentation and a good balance between teaching and research. Researchers might choose safer options and not choose to participate in experimental approaches. Likewise, they might prefer to prioritise teaching rather than research. Finally, there is some evidence that indicators cannot account for the specific patterns of impact and publications in the SSH. However, further studies are needed to shed light on the strategic and behavioural effects of these types of structures being implemented in university systems.

## 6. References

- Aagaard, K., Bloch, C., & Schneider, J. W. (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. *Research Evaluation*, 24(2), 106–117. <https://doi.org/10.1093/reseval/rvv003>
- Aagaard, K., Bloch, C., Schneider, J. W., Henriksen, D., Ryan, T. K., & Lauridsen, P. S. (2014). *Evaluering af den norske publiceringsindikator*. Dansk Center for Forskningsanalyse, Institut for Statskundskab, Aarhus Universitet. [https://pure.au.dk/portal/da/publications/evaluering-af-den-norske-publiceringsindikator\(39dac592-4fd5-4820-8293-53a494b5fed4\).html](https://pure.au.dk/portal/da/publications/evaluering-af-den-norske-publiceringsindikator(39dac592-4fd5-4820-8293-53a494b5fed4).html)
- Adler, N. J., & Harzing, A.-W. (2009). When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. *Academy of Management Learning & Education*, 8(1), 72–95. <https://doi.org/10.5465/amle.2009.37012181>
- Ancaiani, A., Anfossi, A. F., Barbara, A., Benedetto, S., Blasi, B., Carletti, V., Cicero, T., Ciolfi, A., Costa, F., Colizza, G., Costantini, M., di Cristina, F., Ferrara, A., Lacatena, R. M., Malgarini, M., Mazzotta, I., Nappi, C. A., Romagnosi, S., & Sileoni, S. (2015). Evaluating scientific research in Italy: The 2004–10 research evaluation exercise. *Research Evaluation*, 24(3), 242–255. <https://doi.org/10.1093/reseval/rvv008>
- Auranen, O., & Nieminen, M. (2010). University research funding and publication performance—An international comparison. *Research Policy*, 39(6), 822–834. <https://doi.org/10.1016/j.respol.2010.03.003>
- Ávila-Robinson, A., Mejia, C., & Sengoku, S. (2021). Are bibliometric measures consistent with scientists' perceptions? The case of interdisciplinarity in research. *Scientometrics*, 126(9), 7477–7502. <https://doi.org/10.1007/s11192-021-04048-0>
- Bonaccorsi, A. (2020). Two decades of research assessment in Italy. Addressing the criticisms. *Scholarly Assessment Reports*, 2(1). <https://doi.org/10.29024/sar.28>
- Cañibano, C., Vilardell, I., Corona, C., & Benito-Amat, C. (2018). The evaluation of research excellence and the dynamics of knowledge production in the humanities: The case of history in Spain. *Science and Public Policy*, 45(6), 775–789. <https://doi.org/10.1093/SCIPOL/SCY025>
- Cattaneo, M., Meoli, M., & Signori, A. (2016). Performance-based funding and university research productivity: The moderating effect of university legitimacy. *Journal of Technology Transfer*, 41(1), 85–104. <https://doi.org/10.1007/s10961-014-9379-2>
- Coate, K., Barnett, R., & Williams, G. (2001). Relationships Between Teaching and Research in Higher Education in England. *Higher Education Quarterly*, 55(2), 158–174. <https://doi.org/10.1111/1468-2273.00180>
- Dahler-Larsen, P. (2011). *The Evaluation Society*. <https://www-degruyter-com.ep.fjernadgang.kb.dk/document/doi/10.1515/9780804778121/html>
- De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*,

25(2), 161–169. <https://doi.org/10.1093/reseval/rvv038>

De Witte, K., Rogge, N., Cherchye, L., & Van Puyenbroeck, T. (2013). Economies of scope in research and teaching: A non-parametric investigation. *Omega (United Kingdom)*, 41(2), 305–314. <https://doi.org/10.1016/j.omega.2012.04.002>

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New Public Management Is Dead—Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, 16(3), 467–494. <https://doi.org/10.1093/jopart/mui057>

Geuna, A., & Martin, B. R. (2003). University Research Evaluation and Funding: An International Comparison. *Minerva*, 41(4), 277–304. <https://doi.org/10.1023/B:MINE.0000005155.70870.bd>

Grant, J., Brutscher, P.-B., Kirk, S. E., Butler, L., & Wooding, S. (2010). Capturing Research Impacts: A Review of International Practice. Documented Briefing. In *RAND Corporation*. RAND Corporation.

Hamann, J. (2016). The visible hand of research performance assessment. *Higher Education*, 72(6), 761–779. <https://doi.org/10.1007/s10734-015-9974-7>

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. <https://doi.org/10.1093/reseval/rvu029>

Hattie, J., & Marsh, H. W. (1996). The Relationship between Research and Teaching: A Meta-Analysis. *Review of Educational Research*, 66(4), 507–542. <https://doi.org/10.2307/1170652>

Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251–261. <https://doi.org/10.1016/j.respol.2011.09.007>

Hicks, D., Tomizawa, H., Saitoh, Y., & Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation*, 13(2), 76–86. <https://doi.org/10.3152/147154404781776446>

Higher Funding Council of England England. (n.d.). *About the REF - REF 2021 (Worldwide)*. Higher Education Funding Council for England. Retrieved October 26, 2022, from <https://www.ref.ac.uk/about-the-ref/>

Italian National Agency for the Evaluation of Universities and Research Institutes. (n.d.). *VQR – ANVUR – Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca*. Retrieved October 28, 2022, from <https://www.anvur.it/en/activities/vqr/>

Jonkers, K., & Zacharewicz, T. (2016, April 14). *Research Performance Based Funding Systems: A Comparative Assessment*. JRC Publications Repository. <https://doi.org/10.2760/70120>

Kelman, S., & Friedman, J. N. (2009). Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-

Time Target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19(4), 917–946. <https://doi.org/10.1093/jopart/mun028>

Kosten, J. (2016). A classification of the use of research indicators. *Scientometrics*, 108, 457–464. <https://doi.org/10.1007/s11192-016-1904-7>

Lawrence, P. A. (2008). Lost in publication: How measurement harms science. *Ethics in Science and Environmental Politics*, 8(1), 9–11. <https://doi.org/10.3354/esep00079>

Liefner, I. (2003). Funding, resource allocation, and performance in higher education systems. *Higher Education*, 46(4), 469–489. <https://doi.org/10.1023/A:1027381906977>

Mägi, E., & Beerkens, M. (2016). Linking research and teaching: Are research-active staff members different teachers? *Higher Education*, 72(2), 241–258. <https://doi.org/10.1007/s10734-015-9951-1>

Mathies, C., Kivistö, J., & Birnbaum, M. (2020). Following the money? Performance-based funding and the changing publication patterns of Finnish academics. *Higher Education*, 79(1), 21–37. <https://doi.org/10.1007/s10734-019-00394-4>

Melchiorson, P. M. (2019). Bibliometric differences – a case study in bibliometric evaluation across SSH and STEM. *Journal of Documentation*, 75(2), 366–378. <https://doi.org/10.1108/JD-07-2018-0108>

Ministry of Higher Education and Science. (2019). *BFI rules and regulations*. The Bibliometric Research Indicator. [https://ufm.dk/en/research-and-innovation/statistics-and-analyses/bibliometric-research-indicator/governance?set\\_language=en&cl=en](https://ufm.dk/en/research-and-innovation/statistics-and-analyses/bibliometric-research-indicator/governance?set_language=en&cl=en)

Ministry of Higher Education and Science. (2020). *Kriterier for niveauinddeling af forlag på BFI-listen for forlag* [<https://ufm.dk/forskning-og-innovation/statistik-og-analyser/den-bibliometriske-forskningsindikator/BFIs-regler/kriterier-for-niveauinddeling-af-forlag-pa-bfi-listen-for-forlag>]. Den Bibliometriske Forskningsindikator.

Mouritzen, P. E., Opstrup, N., & Bak Pedersen, P. (2018). *En fremmed kommer til byen: Ti år med den bibliometriske forskningsindikator*. Syddansk Universitetsforlag.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review. *Scientometrics*, 66(1), 81–100. <https://doi.org/10.1007/s11192-006-0007-2>

Neumann, R. (1992). Perceptions of the teaching-research nexus: A framework for analysis. *Higher Education*, 23(2), 159–171. <https://doi.org/10.1007/BF00143643>

Ochsner, M., Hug, S. E., & Daniel, H.-D. (2016). *Research assessment in the humanities: Towards criteria and procedures* (p. 247). <https://doi.org/10.1007/978-3-319-29016-4>

Ochsner, M., Hug, S., & Galleron, I. (2017). The future of research assessment in the humanities: Bottom-up assessment procedures. *Palgrave Communications*, 3(1), 17020. <https://doi.org/10.1057/palcomms.2017.20>

OECD, & Butler, L. (2010). Impacts of performance-based research funding systems: A review

of the concerns and the evidence. In OECD, *Performance-based Funding for Public Research in Tertiary Education Institutions* (pp. 127–165). OECD.  
<https://doi.org/10.1787/9789264094611-7-en>

Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, 107(2), 357–368.  
<https://doi.org/10.1007/s11192-016-1845-1>

Socialdemokratiet, Dansk Socialistisk Folkeparti, Radikale Venstre, Enhedslisten, Det Konservative Folkeparti, Nye Borgerlige, Frie Grønne, Liberal Alliance, Alternativet, & Kristendemokraterne. (2021). *Aftale om basismidler til forskning*.

Wang, J., Lee, Y.-N., & Walsh, J. P. (2018). Funding model and creativity in science: Competitive versus block funding and status contingency effects. *Research Policy*, 47(6), 1070–1083. <https://doi.org/10.1016/j.respol.2018.03.014>

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