RoRI RESEARCH ON RESEARCH INSTITUTE

RoRI Working Paper No.5

‘Excellence’ in the research ecosystem: a literature review

Lisette Jong, Thomas Franssen and Stephen Pinfield

September, 2021

About the authors

Lisette Jong is a Postdoctoral Researcher at the Centre for Science and Technology Studies (CWTS), Leiden University. She works as a Research Associate on the RoRI project 'Transforming Excellence: Definitions, Applications and Alternatives to Excellence in the Evaluation, Management and Measurement of Research.' e.p.d.jong@cwts.leidenuniv.nl; https://cwts.nl/people/lisette-jong; https://orcid.org/0000-0001-8001-9462

Thomas Franssen is a Senior Researcher at the Centre for Science and Technology Studies (CWTS), Leiden University. t.p.franssen@cwts.leidenuniv.nl; https://www.cwts.nl/people/thomasfranssen; https://www.thomasfranssen.net/; https://orcid.org/0000-0002-4852-8862

Stephen Pinfield is Professor of Information Services Management at the University of Sheffield and an Associate Director of the Research on Research Institute (RoRI). s.pinfield@sheffield.ac.uk; https://www.sheffield.ac.uk/is/people/academic/stephen-pinfield; @StephenPinfield; https://orcid.org/0000-0003-4696-764X

For citation: Jong, L., Franssen, T.P., and Pinfield, S. (2021) 'Excellence' in the Research Ecosystem: A Literature Review. RoRI Working Paper No. 5., September 2021. **DOI: 10.6084/m9.figshare.16669834**

Acknowledgements

We want to thank the RoRI partners Australian Research Council (ARC), Austrian Science Fund (FWF), the Canadian Institutes of Health Research (CIHR), Fondazione Telethon, the European Molecular Biology organisation (EMBO), the Michael Smith Foundation for Health Research (MSFHR), the Swiss National Science Fund (SNSF) and the Wellcome Trust for their participation in the Transforming Excellence project. We are in particular grateful for the generous feedback we received from our contact persons at these organisations on an earlier version of this working paper.

RoRI RESEARCH ON RESEARCH INSTITUTE

http://researchonresearch.org

Contents

Highlights

- The notion of ‘excellence’ has become an increasingly important part of the research ecosystem over the last 20 years and has shaped science policy, research funding and evaluation activities. Notions of excellence are mobilized in the context of national evaluation systems, institutional funding programs, grant project funding, Centers of Excellence, and play a role in individual career assessment. While omnipresent in the research ecosystem, there is no consensus on what ‘excellence’ means or how it should be recognized. The wide range of approaches to excellence are enacted through bibliometric indicators and intuitive understandings, alongside many others. Such different notions of excellence co-exist in the research ecosystem and come with constitutive effects that shape research, evaluation and funding practices.

- North-American and Western-European contexts of origin have shaped uses of excellence *and* research on excellence initiatives. While measures of excellence are often presented as, or aspire to be, ‘objective’ or ‘neutral’ and standards and ranking systems characterized as ‘global’ or ‘international,’ the notion of excellence itself is thoroughly shaped by the sociopolitical and historical context of its emergence. These political ties need to be made visible in order to understand how the excellence regime (re-)produces inequalities in the research ecosystem.

- Several authors have argued that there is an imbalance between intended and unintended consequences of competition and concentration of resources leading to hyper-competition, and a wide range of associated undesirable behaviors in research and funding practices. While critiques of the excellence regime are as ubiquitous as the notion itself, alternatives are scarcely formulated. There is evidence of attempts at ‘patching’ some of the negative effects of the excellence regime. More fundamentally, there is now a tendency towards pluralizing or diversifying notions of excellence, for example to include measures around ‘impact’ and research cultures. But underlying assumptions about competition and meritocratic ideals remain largely unquestioned.

- Notions of excellence in research funding organisations have been underexplored in the academic peer reviewed literature. At the same time, these organisations are considered to play a key role in the institutionalization of excellence discourses. This calls for a further empirical exploration of the ways in which notions of excellence are used in research funding organisations, the concerns that arise around these uses and what strategies are developed to mitigate the issues.

1. Introduction

The notion of 'excellence' has become an increasingly important part of the research ecosystem over the last 20 years (OECD 2014). Initiatives like the UK's Research Excellence Framework, Switzerland's Eccellenza grants and Germany's Exzellenzinitiative have meant that excellence now plays a crucial role in science policy, research funding and evaluation activities. The prominence of the notion of 'excellence' does not, however, imply that it carries a single or shared meaning. 'Excellence' rather shows characteristics of a 'boundary object' (Star & Griesemer 1989); it is ambiguous and ill-defined which allows for a range of usages.[1] One of the ways it is more narrowly operationalized is through bibliometric indicators, such as the H-Index, Journal Impact Factors and citation scores but just as often the meaning of excellence remains tacit. For example, in the context of peer review, excellence is often recognized through informal 'instinctive' reactions ('I know it when I see it').

Some critics of 'excellence' have called for the concept to be widened or 'pluralized' – drawing on a wider range of criteria and associated indicators, such as societal impact or levels of openness. In other cases, calls have been made to scrap the concept entirely – it being seen as a 'hoorah' word (O'Connor et al. 2020), performing a rhetorical function but lacking any intrinsic meaning. Critics emphasize how the excellence regime creates perverse incentives in the research ecosystem and drives undesirable behaviors, such as 'impact-factor chasing', 'salami slicing' and 'goal displacement'.[2]

We aim to provide an overview of the English language literature published after 2000 on 'excellence' in the academic research ecosystem. This literature review is part of a larger empirical study on notions of excellence in research funding organisations. Hence, the publications ultimately included in this review were selected for their relevance for discussing 'excellence' in relation to research funding activities. The literature review and the empirical study together aim to shed light on how notions of excellence have been mobilized in the context of research funding activities, concerns about the different uses of excellence and the strategies to

[1] 'Boundary objects are both adaptable to different viewpoints and robust enough to maintain identity across them' (Star & Griesemer 1989:387). Star and Griesemer (1989) argue that boundary objects play a crucial role in enabling cooperation between actors from different fields that relate to the same object in different ways. 'Excellence' could be considered as a boundary object that plays a role in establishing cooperative relations between actors in the fields of research, research funding, science policy and higher education. Albeit the 'robustness' of the notion of excellence remains a point of discussion.

[2] Impact factor chasing is a term used to address a strong focus on publishing in high impact factor journals, possibly at the cost of other activities. Salami slicing refers to cutting up the research in the smallest publishable parts to increase the total number of publications from a study. Goal displacement happens when scoring high in an evaluation or ranking system becomes a goal in itself instead of a means to evaluate performance. These behaviors can be considered examples of 'gamesmanship' that is encouraged by an emphasis on 'excellence' (Moore et al. 2017).

mitigate these. In doing so, we hope to inform the debate around excellence, competition and meritocratic ideals in the research ecosystem.

1.1 Method

To assemble an initial corpus of potentially relevant literature, we searched the Web of Science database for 'excellence' as topic term ('excellence' appearing in title, abstract and/or keywords of a publication) in a selected list of ten journals in the fields of science policy and evaluation studies, science and technology studies (STS) and scientometrics (*Research Policy*, *Science and Public Policy, Minerva, Social Studies of Science, Science Technology & Human Values, Research Evaluation, Scientometrics, Higher Education, Higher Education Policy, Science as Culture*) where we expected to find publications on excellence in the research ecosystem. This resulted in an initial sample of 279 articles.

To broaden our initial sample, we used a citation link report produced with the CWTS in-house version of Web of Science.[3] This citation link report listed 3397 additional publications that cite one or more of the publications in our initial sample. We then narrowed this list down to the 114 publications that had at least four links to our initial sample and/or have the term 'excellence' in the title. Further literature was added by looking through the reference lists of key publications. In this way we also incorporated 'grey literature', such as reports, and books into our analysis. Sharing an earlier version of the literature review with the strategic partner organisations in the empirical research project resulted in generous feedback and additional resources on a variety of aspects of excellence, including (locally) relevant issues around equity, diversity and inclusion and Indigenous rights that did not immediately come to the fore in the academic literature.

Because of our focus on research funding activities we excluded publications that were solely focused on excellence in the context of education. We are aware that searching with 'excellence' as a topic term may have excluded potentially relevant publications where for example 'research quality' or other terms are used in a synonymous way with notions of excellence. Another limitation is our exclusive focus on English language publications. While the literature on excellence initiatives is already quite centered on the European context, these constraints and the fact that all three authors are based in Europe, have likely steered our initial sample in that direction as well.

In our thematic discussion of 'excellence' below, we begin with a brief overview of the emergence of the notion of excellence in the scholarly ecosystem. We also discuss different conceptualizations of excellence as an object of study in STS, science policy and evaluation studies. We go on in the second section of this paper to focus on different aspects of the

[3] We want to thank our colleague at CWTS, Ludo Waltman, for his generous help with producing the citation link report.

academic research ecosystem as these relate to funding activities, such as research policy and evaluation, funding arrangements and peer review. In addition, we address how excellence-driven funding programs shape research institutions, projects and the everyday work of researchers. From this body of literature, we distil recurring themes in the debate about the 'constitutive effects' (Dahler-Larsen 2014) of the different uses of excellence in the research ecosystem. We elaborate on these in the third section of this paper. In the fourth section, we outline strategies designed to counter the negative effects associated with the emphasis on excellence in the research ecosystem. These comprise 'patching' of existing uses of excellence, 'pluralizing' excellence, and attempts to move beyond excellence by radically 'transforming' the ways in which research activities are organized and valued.

1.2 The emergence of research excellence

Flink and Peter (2018) trace 'excellence' from its use in 1950's US Cold War discourse in which it figured as a term to encourage what was seen as tendency towards mediocracy, to excellence becoming a 'euphemizer' for competition in the context of European science policy. Along the way, excellence shifted from a qualifier of individual performance to a stratification device applied at the level of research groups, universities and even countries.

However, the notion of excellence in research did not develop in a vacuum. From the 1980s onwards, excellence became an important concept in many sectors. Most notably, in the area of business management, the notion of excellence mobilized a neo-liberal focus on self-organizing and investment in human capital as precursors for economic growth (Flink & Peter 2018) and has given rise to a plethora of management excellence models (Dahlgaard-Park & Dahlgaard 2007). In the research and higher education sector, the excellence discourse was driven by meritocratic ideals, a need to legitimize public spending on research and the increasing competition for scarce resources (Münch, 2015). One of the most notable effects of the excellence discourse is the increasing importance of university rankings that shapes the global field of research and higher education, despite sustained criticism (Bougnol & Dulá, 2015; Brankovic et al., 2018; Johnes, 2018).

Hammarfelt et al. (2015) date the first university ranking back to 1910, produced by the American psychologist James McKeen Cattell. Informed by eugenicist concerns about a perceived decline in 'eminent men,' Cattell focused his studies on 'men of exceptional ability.' He produced a biographical directory listing American scientists[4] and calculated university rankings by measuring 'productivity' (expressed as the number of 'scientific men' in one location) and

[4] Studies attempting to grasp the personality traits and abilities of excellent researchers are not just something from the past. See Aráujo et al. (2017) for an example of a more recent psychological study with 'six of the most prominent Portuguese scientists' that approaches excellence as a quality of individuals.

'performance' (as judged by peers active in the same discipline) (Godin 2007). Building on the work of the 19th century statistician Francis Galton[5], Cattel laid the grounds for the use of bibliometric indicators and peer review to evaluate the performance of individual scientists and research institutions. Cattel argued that these rankings could be used by students in choosing a university, for international comparison and for raising the status and salary of researchers through competition between institutions. He also suggested a salary scale based on the ranks of individual scientists (Hammarfelt et al. 2015). The constitutive effects of Cattel's efforts to make 'eminence' measurable resonate with contemporary uses of excellence indicators:

> *In fact, the act of measuring eminence defined what would henceforth count as eminence. As a result, the concept lost the ambiguity it could still retain when it was not yet stabilized by the process of measurement. Comparable processes of definition and social construction can now be seen around contemporary enactments of excellence and quality.* (Hammarfelt et al. 2015: 393)

The contemporary emphasis on excellence has been associated with a process of alignment of quality notions in the field of higher education and research globally (Brankovic et al., 2018; Langfeldt et al. 2020). This 'global script' of excellence specifically favors the model of the research intensive university as it emerged in North-America from the 1950s (Antonowicz et al. 2017). It prescribes a model university focused on reputation, measurable outcomes and the global competition for scarce resources. Widespread indicators, monitoring and ranking techniques aimed at identifying and measuring excellence, originate predominantly in Western European and North American countries and are geared to this specific mode of knowledge production. The link between performance measures and the model of the research intensive university limits the space for the development of epistemic practices that do not result in measurable outcomes.

In the 1990s, the notion of excellence entered more clearly into the European policy context. In 2000 the European Union launched the Lisbon Strategy aimed at making Europe 'the most competitive and dynamic knowledge-based economy in the world'.[6] The notion of excellence was mobilized as part of this strategy to improve Europe's competitive position in what came to be understood as a 'global' knowledge economy (Flink & Peter 2018). In the Lisbon Strategy, excellence was used as a general concept, broadly referring to the quality of research practice and aimed at fostering knowledge production. The European Union's 2013 'Composite Indicator for Scientific and Technological Research Excellence,' marked a shift towards a more narrow definition of excellence focused on measurable outputs and making performance comparable across countries, institutions and research groups. Following Sørensen et al. (2016) it was the

[5] Having coined the term in 1883, Galton is also known as the 'father of eugenics.'
[6] https://portal.cor.europa.eu/europe2020/Profiles/Pages/TheLisbonStrategyinshort.aspx

'increasing pressures of global competition, tightening resource constraints that were further exacerbated by the financial crisis together with a growing emphasis on auditing and performance measurement in the public sector,' that informed the development of the indicator and the reductionist approach to excellence that it advanced. The development and implementation of such measures were thus motivated by a desire to promote European research informed by an understanding of the research ecosystem as a competitive space in the 'global' knowledge economy.

Importantly, while measures of excellence have often been presented as, or aspire to be, 'objective' or 'neutral' and standards and ranking systems characterized as 'global' or 'international,' the notion of excellence itself is thoroughly shaped by the sociopolitical and historical context of its emergence. This becomes apparent when an excellence model is explicitly adopted for institutions to encourage a move towards the model of the research intensive, competitive university that prevails in Western Europe and North America, and specifically when frictions appear in translating the excellence script into different local contexts. A case in point is the introduction of 'excellence' since the mid-2000s in research policies of Central and Eastern European Countries as part of the process of 'Europeanization after Communism' (Radosevic & Lepori 2009). The adoption and spread of the excellence agenda in post-colonial settings has also been considered as both a consequence of the historical development of academic institutions under colonial rule, and as a new form of epistemic colonization that reinforces existing inequalities in the research ecosystem (Neylon 2020). This, and the above-mentioned historically gendered nature of the notion of excellence, will be further addressed later in this paper.

1.3 Scholarly reflections on the notion of excellence

In the academic literature there seems to be agreement that excellence is a fuzzy notion, a rather ambiguous term to which different meanings can be ascribed. The reluctance to define excellence was addressed by Merton as early as the 1960s:

> *Many of us are persuaded that we know what we mean by excellence and would prefer not to be asked to explain. We act as though we believe that close inspection of the idea of excellence will cause it to dissolve into nothing....* (Merton 1973: 422-423)

In his essay 'Recognition and Excellence: Instructive Ambiguities' originally published in 1960, Merton differentiated between excellence as a quality of something or someone, and excellence as performance. He argued that excellence as an attribute or characteristic inherent to

someone—a potential—can be measured or accounted for.[7] Excellence as performance refers to someone's achievements. Merton mobilized this conceptualization to point at the gap between capacity and performance at both the individual and group level and to address inequalities in the opportunities to develop potential. Thus in his efforts to define the notion, Merton attached excellence to his meritocratic ideals: for excellence to be encouraged, meritocracy should be fostered. This meritocratic understanding of excellence has been influential in legitimizing its role in higher education and research (Flink & Peter 2018).

In more recent literature, the nebulous nature of 'excellence' is also discussed. Within this literature an important difference is the extent to which this fuzziness of excellence comes out as either its core problem or strength. Excellence has been characterized as an 'essentially contested concept' (Ferreti et al. 2018), demonstrated by the controversies that rise when excellence is operationalized and quantified in the context of indicator development. Ferreti and colleagues argue that difficulties in agreeing on a shared meaning of excellence, is an effect of the notion's inherent openness since excellence is 'of no physical substance itself' (2018: 732).

The lack of consensus around the meaning of excellence has alternatively been seen as a positive quality. In this sense, the fuzzy notion of excellence is understood to operate as a 'boundary object' (Star and Griesemer 1989) that enables actors from different fields, working with different understandings of excellence, to cooperate around it. Excellence seems to figure as a 'boundary object' in research funding organisations as they operate between science policy and research communities (Langfeldt et al. 2020: 125), with excellence creating a space where political and academic values and priorities can coexist (Hellstrom 2011: 118).

In the context of her study of interdisciplinary peer review panels, Lamont conceptualizes excellence as a 'quintessential polymorphic term' (2009: 159). Panelists come with a range of different meanings of excellence to assess an equally diverse range of proposals. The panel's model of deliberation allows for flexibility in the application and weighing of criteria in relation to individual proposals and the shared standards that emerge in such discussions are temporary and contextually informed. This room for expert judgment and flexible uses of the notion of excellence is, according to Lamont, the added value of peer review panels over 'mechanized forms of evaluation that attempt to remove all traces of the personal, the idiosyncratic and the inconsistent' (2009:51).

However, the room for manoeuvre that comes with an ambiguous notion like excellence, has also been seen to have less desirable effects. O'Connor et al. (2020) argue that the 'vagueness' and perceived neutrality of the notion of excellence facilitate its use as a 'rationalizing myth' that

[7] Merton was however critical of the use of universal tests and measurements to identify potential. Instead, he suggested the employment of local 'talent scouts' to recognize excellence in their respective communities (1973: 427).

obscures gender-biased practices in the evaluation of candidates for faculty appointments.[8] Moore et al. (2017) draw on the aforementioned work by Lamont to conceptualize excellence as a 'linguistic interchange mechanism,' but contrary to Lamont, denounce its usefulness as such by rooting problems with reproducibility, fraud, conservatism and hyper-competition in the 'excellence rhetoric.'

In summary, in contemporary literature excellence is conceptualized as a discursive construct in order to shed light on its performative effects in specific practices in the research ecosystem. Thus the notion of excellence is understood as a word that can come to mean different things in different situations to different effects and purposes. However, in most of the empirical literature about research excellence, the notion of excellence itself has come under less scrutiny. This literature includes efforts to define and quantify excellence through (bibliometric) indicators. It also includes empirical research articles that focus on excellence initiatives for example, the distribution of funding and knowledge production practices, or how operationalizations of excellence function in evaluation practices. The following sections engage with this literature.

[8] While the study by O'Connor et al. specifically focused on gender as a marker of difference, Mohamad & Beagon (2019) show how racialized and Indigenous faculty also are disadvantaged by 'the myth of colour-blind meritocracy.'

2. Excellence in research funding practices

Excellence is omnipresent in the research ecosystem. The term is used in relation to funding programs, evaluation scales and mission statements. It is applied at the level of individual researchers, research groups, institutions and countries in activities of governing, funding, evaluating and performing scientific research and output. While the focus in this section is on research funding organisations, the activities that shape funding in the research ecosystem stretch well beyond organisations that allocate resources on a competitive basis. These organisations operate in the context of science policies promoting excellence, and the effects of the focus on excellence in the wider science system can, for example, be observed in everyday decision making in research groups, the priorities set by research institutions and the formation of collaborations.

The following discussion, therefore, includes empirical studies that attend to excellence as it figures in science policy and evaluation, the development of funding schemes, allocating grants and the effects of funding. In mapping this body of literature we found that the work of funding organisations, despite playing a key role in the science system, is sparsely covered in peer-reviewed publications. The literature that addresses excellence in the research ecosystem targets either the development of science policies and research evaluation practices, or the effects of excellence-based science policies on institutions, research groups and individual researchers. The role of 'excellence' in research funding organisations and the activities they manage is often not covered.

Before discussing this literature, we focus on the quantification of research excellence through bibliometric indicators that have thoroughly shaped the research ecosystem in the past decades.[9] One key feature of the current research ecosystem is how a system of science in which quality notions are specific to knowledge communities and enacted through expert peer review, has come to co-exist with quantification based techniques for assessment aimed at ranking and comparison across those communities (Langfeldt et al. 2020). Rather than taking these assessment practices for granted, we aim to explore how performance measures came about and the role that the development of metrics has played in shifting understandings of excellence.

[9] Although the emphasis on bibliometrics affects every field of science it does not affect all disciplines in the same way. For example, publication practices in the humanities differ from those in the natural sciences. The latter field is better covered in databases used for citation analyses than the former. This led to some national evaluation systems differentiating their assessment practices with units in the natural sciences being evaluated based on bibliometric indicators and units in the social science and humanities based on peer review (Aksnes et al. 2019).

2.1 Making excellence measurable

The Science Citation Index, accessible through the Web of Science, was first published in 1964 and originally developed with information dissemination and retrieval purposes in mind. However, its tracking of citations also allowed for quantitative analyses of scientific output (Wouters 1999). The field of scientometrics got a boost in the 1970s and especially 1980s due to the increasing interests of governments in justifying investment in higher education and research. As an academic field, the establishment of its own journal *Scientometrics,* launched in 1978, was an important moment (Garfield 2007). Initially, discussions in scientometrics, then tightly connected to another emerging field: science and technology studies (STS), centered around exploring the possible uses of this new data source, not just for science policy but for the sociology of science as well (Franssen & Wouters, 2019). The issue of how citations relate to quality was an open question.

From the 1990s onwards, scientometric scholarship grew away from STS and focused on methodological issues around the improvement of indicators or developing new ones and the debate about what citations scores measure abated somewhat (Aksnes et al. 2019). The main benefit of quantitative approaches to research excellence is that they apparently enable easy and seemingly clear comparison between researchers, research groups and universities. Simple measures of excellence are attractive in a science system which is characterized by finite funding and, consequently, the need to make decisions about resource distribution. As Ferretti et al. put it: 'Indicators of research excellence are both hard to live with and without' (2018:738-9).

In an attempt to make excellence measurable, Tijssen (2003) emphasizes that excellence is a 'comparative measure' referring to 'surpassing something or someone in some quality.' Rather than precisely pinning down excellence, Tijssen proposes a 'fuzzy multi-dimensional model' that addresses a set of interrelated dimensions and combines bibliometric indicators and expert judgment, while no single dimension is necessarily essential for grasping excellence entirely (2003: 94). This interactive scoreboard of research excellence is not explicitly intended as a stand-alone evaluation device but rather provides an empirical framework for expert-based informed debate. Following Tijssen, the scoreboard should be adjusted to each organisational context and decisions should not solely be based on the data.

The complexity that Tijssen (2003) attempts to account for gets lost in quantitative approaches that simply define excellence as referring to the upper tail of a performance distribution, with the cutoff point for what qualifies as excellent most commonly set at the top 1, 5 or 10% (Bonnacorsi et al. 2017; Tijssen & Kraemer-Mbula 2018). For example, scientometricians addressing the 'transatlantic gap,' understood as the measured distance between Europe and the US in terms of

research excellence, provide their definition as follows: 'We defined scientific excellence in terms of the top 10% of the world distribution of separate and transparent indicators, based on the volume of publications, the number of citations, and the share of publications in the top 10% SNIP journals' (Bonnacorsi et al. 2017:238). Such an emphasis on journal publications covered by bibliographic databases in defining research excellence disadvantages researchers and institutions that, for example, find their work less well covered in those databases, including those that publish in languages other than English.

Metrics are often claimed to be more efficient and objective in comparison to more time and cost intensive ways of quality assessment, like peer review (e.g. Abramo et al. 2011). However, metrics are not neutral. The process of quantification actively shapes what comes to count as excellent performance and comes with very specific and often narrow definitions of excellence. As several authors point out, there is a tendency to value what can 'easily' be measured (Tijssen 2020; Sørensen et al. 2016). As such, the quantification of research excellence is as much a political and normative issue, driven by trends in evidence-based policy making (Ferretti et al. 2018).

How quantification sets limits to what may be eligible for inclusion in a metrics-based approach to excellence is exemplified by the basic criteria of observability that need to be met for outputs, impacts or research efforts to be perceived as 'excellent', as identified by Tijssen and Kraemer-Mbula (2018). These are the need to be : 1) visible and recognizable (to others); 2) attributable (to research contributors and participants); 3) comparable (within a generally accepted frame of reference); 4) categorized in terms of quality judgement (by external experts or other observers) (2018: 394). As a consequence, potentially meaningful outputs that do not meet such criteria of observability tend to be excluded when a metrics-based approach to excellence is applied.

With the spread of metrics based approaches of quality assessment, several authors warn that narrow quantitative understandings of excellence can come to be used synonymously with 'research quality' (e.g. Vessuri at al. 2013). Aksnes et al. (2019) more specifically suggest that citation metrics are not suitable for evaluating the plausibility, originality and societal value of research, but can function as proxies for scientific relevance and scientific impact, rather than as a direct indicator of quality. An emphasis on performance metrics in assessments has been associated with goal displacement and 'gaming' behavior, a bias against interdisciplinarity, a devaluation of other academic activities,[10] a push towards mainstream topics, and use of institution level indicators for assessing individual researchers (De Rijcke et al. 2016). It is also in this regard that widely supported calls for responsible use of metrics in research funding and evaluation practices have emerged in the last decade. The Declaration On Research Assessment

[10] Activities like teaching and service work tend to be taken up more often by women and other historically excluded groups such as racialized and Indigenous academics (e.g. Thornton 2013; Mohamad & Beagon 2019).

(DORA, 2012) urges actors in the research ecosystem to not use Journal Impact Factors as a quality measure in any form of assessment. The Metric Tide report developed the idea of 'responsible metrics' characterized by robustness, humility, transparency, diversity, and reflexivity (Wilsdon et al., 2015). The Leiden Manifesto (Hicks et al. 2015) provides ten principles for responsible use of metrics in evaluation and decision making processes.

2.2 Excellence-based funding: national performance-based research funding systems

The way research funding is allocated has changed greatly over the last decades. Consequently, research on research funding is of increasing interest in science policy studies (Gläser & Velarde, 2018). One of the important changes in the allocation of research funding is the rise of performance-based funding schemes (Auranen & Nieminen, 2010; Thomas et al., 2020), often tied to a national research evaluation process. Another is the rise of competitive project funding programs for individuals or research teams (Franssen & De Rijcke, 2019; Bloch & Sørensen, 2015).

One driver behind excellence policies to manage the science system on the national level is the competition for scarce resources and accountability for the public money spent on research. The objective to 'identify excellence' translates into evaluation systems that rank the performance of research and higher education institutions against (inter)national benchmarks. The Research Assessment Exercise (RAE) in the UK, launched in 1986, was the first of many Performance-based Research Funding Systems (PRFSs) to emerge as national systems that govern the distribution of research funds based on ex-post evaluation of research outputs (Hicks 2012; Thomas et al. 2020). Not all national evaluation systems directly steer the distribution of funding, or have over time shed this function, as is the case with the Australian initiative, Excellence in Research for Australia (ERA). These systems are mainly considered to hold relevance in terms of prestige and accountability.

PRFSs generally rely on bibliometric analyses of outputs and peer review, either in combination or as discipline specific measures of quality. The successor of the UK RAE, the Research Excellence Framework (REF), has received considerable attention from scholars in the field of science policy and evaluation studies (Thomas et al. 2020: 278; Sivertsen, 2017). The REF significantly remains a process based primarily on peer review rather than bibliometric indicators. Next to discussion about the processes of the REF, its criteria – particularly the inclusion of 'impact'[11] (Watermeyer, 2016) – and funding formula (Pinar 2020), the increasing costs of the REF have raised questions about the legitimacy of such a resource intensive exercise (e.g. Martin

[11] Further exploring the concept of impact is beyond the scope of this review. But not unlike excellence, it is an ambiguous term that assumes different meanings in different practices.

2011). In this regard, Balbuena (2018) suggests using machine learning (next to peer review) to save costs that can then be invested in the higher education institutions that need it. The Norwegian PRFS is similarly well studied (Schneider 2009; Sivertsen 2016), including research on its effects on publication practices (e.g. Aagaard, Bloch & Schneider 2015).

While elements of PRFSs are adopted between countries, PRFSs do not develop in a linear direction. For instance, in Sweden a model similar to the UK REF was designed but never implemented (Sivertsen 2017) and in Australia direct distribution of funding is no longer part of the ERA-framework. Also the form of excellence-based funding programs differs between countries (Hou et al. 2012; Cremonini, Horlings & Hessels 2018), some target universities others aim to establish centers of excellence, but excellence-based funding programs are spreading across the globe (Antonowicz et al. 2017). Excellence initiatives such as the Double World-Class project in China (Zhao & You 2019) and the World-Class University Project in Taiwan (Fu et al. 2009) do not include all universities but concentrate on a selection of already well-performing universities aiming to further improve their position in global rankings. Despite their differences, these programs seem to share a commitment to the idea that concentrating resources on a relatively small number of presumably the best researchers or institutions will increase the scientific quality and impact of research. However, as Thomas and colleagues note (2020: 283), there is a lack of comparative analytical research on PRFSs that would shed more light on the differences and similarities between PRFSs in different countries.

2.3 Excellence-based funding arrangements: grant funding programs

Scholten et al. (2021:2) define excellence funding as 'highly competitive funding arrangements through which funding is allocated primarily based on considerations of scientific performance (often operationalized as numbers of publications and citations), with little direct consideration of societal relevance or thematic programming.' They thus separate this type of funding programs from mission-oriented funding programs. The latter, which are tied to explicit societal goals, are also increasingly important (Mazzucato 2018). Nevertheless, mission-oriented funding programs are similarly competitive, and often include similar 'excellence' criteria and can in the same way as excellence-based funding programs contribute to the concentration of funding among a small number of researchers (e.g. Bol et al. 2018; Aagaard, Kladakis & Nielsen 2019).

The literature identifies three types of excellence-based funding programs. The first is project grants and prizes for individual researchers (Franssen et al. 2018). A prominent and highly influential European example are the grants allocated by the European Research Council (ERC) (Luukkonen 2014; Hoenig 2017; König 2017) that according to its mission statement has 'scientific

excellence' as its 'sole criterion for selection.'[12] Laudel and Gläser (2014) have studied the epistemic properties of ERC-funded research and conclude that its particular characteristics, notably the long time horizon, flexible and large budget, and risk-tolerant selection process, enable innovative research. Other studies of the ERC include studies of the peer-review process (Luukkonen 2012; van den Besselaar, Sandström & Schiffbaenker 2018) and the organisational responses of universities to the establishment of the ERC (Cruz-Castro, Benitez-Amado & Sanz-Menéndez 2016).

A second type of excellence-based funding arrangement are Centre of Excellence grants. Centre of Excellence grants are characterized by a long time horizon and large budgets and are awarded through a competition (Langfeldt et al. 2015). The aim of such grants is to concentrate resources and create critical mass (Bloch & Sörensen 2015) around specific areas of research in which scientific performance is then enhanced. Centre of Excellence initiatives can be found in different national science systems including Canada (Fisher, Atkinson-Grosjean & House 2001) the Nordic countries (e.g. Langfeldt et al. 2015; Borlaug 2016; Borlaug & Langfeldt 2020; Hellström 2011; Hellström Jabrane & Brattström 2018a; Hellstrom 2018b) and Germany (Schröder et al. 2014). These studies are focused on the role of centers of excellence in the science system as well as their organisational characteristics and effects on epistemic practices.

Lastly, in some countries a selected number of universities receive additional resources through excellence-based funding programs that aim to support the development of elite research-intensive universities. The German excellence initiative, that includes a variety of excellence-based funding programs, is a prominent European example (Schröder et al. 2014; Kehm 2013). Similar policies have also been implemented in, amongst others, Russia (Lovakov et al. 2021), China (Zong & Zhang 2019; Zhao & You 2019), Taiwan (Fu, Baker & Zhang 2020) and Japan (Yonezawa & Shimmi 2016). This literature describes the role of science policies in the national science system and often aims to understand its effects on scientific performance in particular as it affects the standing on university rankings.

These different excellence-based funding arrangements demonstrate that the excellence discourse is widespread in research funding allocation for individual researchers, research groups and consortia as well as institutions. There are, however, large differences between national science systems in both the type of funding arrangements used and the properties of those funding arrangements that deserve further exploration.

[12] See ERC mission statement: https://erc.europa.eu/about-erc/mission.

2.4 Allocation of excellence funding through peer review

In the activities of funding organisations tensions may become visible between policy-oriented notions of excellence, on the one hand, and varying notions of excellence held by peer reviewers embedded in their respective disciplines, on the other hand (Langfeldt et al. 2020). Peer review has been considered the gold standard for identifying excellence. However, notions of excellence in peer review panels are not stable, but rather situationally defined in a process of deliberation (Lamont 2009). Van Arensbergen et al. (2014) shed light on peer review panel decision making practices in a literature review that, in addition to including Sociology of Science and Science Policy Studies, purposefully includes literature from social psychology, to frame peer review as a process of social interaction in which 'non-quality-related factors' contribute to the unpredictability of panel peer review outcomes. In competitive funding programs that claim to identify scientific excellence solely based on merit, factors that play a role in grant allocation, funding decisions and talent selection include alumni relations (Jang et al. 2017), specificities of the social and organisational processes of peer review (Langfeldt 2006, Van Arensbergen et al. 2014) and gender bias (Tamblyn et al. 2018). Several research funding organisations attempt to mitigate these issues by, for example, by providing bias training for their reviewers.[13]

The question has been raised whether review panels actually manage to select the 'best of the best' through ex-ante evaluation. Empirical studies comparing career paths and performance of accepted and rejected applicants of competitive funding schemes find no differences or only slightly better results for researchers in the accepted group (e.g. Hornborstel et al. 2009; Schneider & Van Leeuwen 2014; Van den Besselaar & Sandström 2015). The question whether funds are allocated to the best of the best in excellence schemes has also been addressed in research. In a qualitative empirical study of the ERC peer review process, Luukkonen (2012) observed that reviewer's value judgments are based on current knowledge paradigms, which puts unconventional proposals at a disadvantage. In addition, reviewers aim to avoid 'unjustifiable risk-taking' in funding decisions, which translates into more conservative decision making overall. According to Luukkonen (2012), these tendencies are a constraint to the ERC's mission to fund groundbreaking, excellent research.

Hicks and Katz (2011:143) draw attention to what they call the 'tension between excellence and equity' shown in the way research performance and merit-based research funding decision making takes place. Drawing on empirical data from the UK RAE, they show how research

[13] See for example Canadian Institutes of Health Research training modules on unconscious bias and sex and gender considerations in peer review (https://cihr-irsc.gc.ca/e/50559.html) and the fact sheets on implicit bias by the National Institutes of Health (https://diversity.nih.gov/sociocultural-factors/implicit-bias). Amongst others, the Australian Research Council encourages reviewers to take 'Implicit Associations Tests' (https://www.arc.gov.au/peer-review/process).

performance, measured at the level of university departments in terms of output, is extremely unequally distributed. The authors point out that 'any resource distribution that truly mirrored the distribution of research performance would be untenable in the public realm because of its extreme degree of inequality' (Hicks & Katz 2011:148). Hicks and Katz conclude that funding decisions based on peer review, therefore, do not reflect the distribution of research performance and argue that such a 'mismatch' suppresses incentives for excellent researchers to reach for the top, in other words, concerns about equity lead to compromises in excellence. The paper by Hicks and Katz (2011) adds to the abovementioned work in highlighting that excellence funding is in practice not allocated based on 'merit'[14] alone, neither granted solely to the best of the best. Hence, the everyday reality of excellence programs problematizes the meritocratic ideals that underpin such funding arrangements (see also Scholten et al. 2018).

2.5 Effects of excellence-based funding arrangements on research practices

Several empirical (qualitative) studies have focused on how the conditions of excellence funding affect the work of individual researchers and research groups in terms of epistemic practices, the organisation of research and interdisciplinary relations. Laudel and Gläser (2014) explored the links between the epistemic properties of the research process and the conditions of funding for ERC grant holders. Interviewees stated that there was often no equivalent funding scheme to which they could apply for their envisioned activities. The duration (5 years) and the sum of the ERC grant provided them with the resources to, for example, establish a stable team and buy expensive equipment. Interviewees suggested that the ERC awarded grants to projects that were by other funders deemed 'too risky.'[15]

Excellence-based funding programs like the ERC grants, build on a model for small-scale collaborative experimental science that is common in the biosciences but not as widespread in the social sciences and humanities, where the organisation of research is more individualistic. Laudel and Gläser (2014) expect the ERC model to promote collaborative interdisciplinary research in the social sciences and humanities. In Centers of Excellence schemes, interdisciplinarity is particularly valued and even expected (Hellström et al. 2018). In humanities

[14] However 'merit' may be defined. Hicks and Katz draw on 'output' as defined in the UK RAE, as a performance indicator.

[15] Thus in the context of the Laudel and Gläser (2014) study, the ERC programs were positively associated with risk taking. In Luukkonen's (2012) study, discussed here earlier, the peer review process that guides ERC funding decisions was considered to induce risk avoidant decision making. This apparent tension could be understood in light of the different perspectives taken in the respective studies. Researchers perceive the ERC as relatively open to risky projects compared to other funders, while from within the peer review panels risk avoidance strategies can be observed. As such the two findings do not necessarily contradict one another, but together shed light on the complexities and ambiguities around notions of excellence and perceptions of excellence initiatives.

departments in Norway that hosted a center of excellence, Borlaug and Langfeldt (2020) indeed found an increase in collaborative activities in the form of larger research projects and the establishment of interdisciplinary relations. As publications remained mainly individualist or within previously established collaborative relations, the overall working pattern did not, however, seem to change as much. The organisation of collaborative activities may have been partly symbolic in order to meet grant requirements. Specific to the humanities centers, compared with other disciplines, was the tension around the perceived 'excellence' of the work. One center leader suggested that what is excellent is more contested in the humanities than in other fields because it is considered to be less clear what the relevance of research in the humanities is (Borlaug & Langfeldt 2020: 1753).

Scholten et al. (2021) compared the organisational and epistemic practices of research groups in the Netherlands that did receive excellence funding with those that did not in the same disciplines. They found that the groups with excellence funding enjoyed greater autonomy in the institutional context across all disciplines. The excellence funding provided these groups with a high level of independence that resulted in more freedom to decide on financial and human resource matters and greater control over their own research agendas (see also OECD 2014). This translated into flexible decision making in the course of the research and space to engage in additional research activities like creating their own datasets or setting up new networks. Such effects were even greater in the context of groups that received prize money (Franssen et al. 2018).

Studying the way of working of four high-performing research groups in the Netherlands and Denmark, Degn et al. (2018) found these groups to show characteristics of communities of practice, including the development of internal criteria for recognizing excellence or talent. Rather than about output measures (e.g. grades and publication metrics), these criteria are about the candidate's desired role in the research team, motivation and drive. Shared notions of teamwork and relative independence of the group from the institutional context, strengthened the group identity of the research teams. The authors address the question of causality: is it the internal organisation that makes the group successful in the excellence regime, or is it because the group is successful in acquiring sufficient resources and freedom that it can develop as a community of practice? The issue of causality has also been raised in the case of high-performing researchers in Swedish centers of excellence, who seem to be more heavily engaged with knowledge diffusion (seen as a marker of excellent performance in Swedish science policy) than their colleagues in more precarious positions (Wigren-Kristoferson et al. 2011). Long-term funding gives researchers space to be involved in dissemination activities they enjoy, while they also experience pressure to hold themselves accountable to society for receiving such a large sum of public funding.

Another effect of excellence funding regimes is what Scholten et al. (2021) call 'strategic anticipation': behavior that attempts to maximize the chances of grant success. Strategic anticipation includes efforts to acquire new grants and making publication decisions with the aim of acquiring excellence funding in mind. Such behavior has been observed across disciplinary groups and both in research groups that were and that were not successful in acquiring excellence funding, although the latter groups experienced more pressure. Cañibano et al. (2018) describe how the focus on output (articles) in the Spanish national research evaluation system similarly affects publication decision making of researchers in a history department in Spain. The incentive to publish in high-impact journals, undermines disciplinary values of originality and innovation.

Another form of strategic anticipation can be seen in relation to the criterion of mobility included in funding programs. Mobility is perceived as a proxy for internationalization, excellence and competitiveness and as such tied to career progression (Ackers 2008). As the fact of mobility has come to count more than the quality of experiences, researchers are driven to move internationally in anticipation of acquiring grants or positions. However, the opportunities to move are not equal. Amongst other minoritized groups in academia, the requirement of international mobility may in particular disadvantage Indigenous researchers who have local priorities and obligations (Staniland et al. 2020). Ackers (2008) calls for a qualitative evaluation of diverse experiences that foster internationalization beyond narrow understandings of mobility.

It is important to note that the excellence discourse is only one aspect of the changing science system that, potentially, influences research and publication practices. The excellence-discourse is part of a web of changes, including the rise of project funding, increasingly short-term contracts and hyper-competition in the academic labour market (Fochler, Felt & Müller, 2016; Franssen & De Rijcke, 2019) as well as the rise of new public management, which meant a rise of metricized evaluation and assessment processes on all levels of the science system (De Rijcke et al., 2016).

3. The excellence debate

The notion of excellence and the meritocratic ideals that it rests upon have often been critiqued for its perverse effects on the research ecosystem. These commentaries range from speculative remarks to arguments based in empirical studies about how notions of excellence redefine and shape research practices. Recurring themes in this debate are the (un)desirability of cumulative advantage (the Matthew effect), the incentives that lead to hyper-competition, the perceived tension between 'global' measures of excellence and local relevance of research, and the way in which what might be called 'the excellence regime' (where cultures and processes are significantly shaped by the notion of excellence) promotes conservatism, and homogeneity with the effect of (re)producing inequalities. These concerns are discussed in the following sections.

3.1 The Matthew effect: cumulative advantage

Merton described the Matthew effect in science as follows: 'the Matthew effect consists of the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark' (1973: 4455-446). At the level of individual scientists, a Matthew effect can be observed in the visibility of outputs: the publications of already recognized scientists attract more attention than those of early career researchers. Merton also addressed cumulative advantage at the institutional level: centers recognized for scientific excellence are allocated larger resources and attract a larger share of talented students. The (un)desirability of cumulative advantage in the research ecosystem is an ongoing matter of debate associated with the excellence regime and competitive funding programs.

One of the arguments in favor of the concentration of funding on the best performing researchers and institutions is the promotion and maintenance of scientific excellence (Bloch & Sörensen 2015; Aagaard et al. 2020). Concentrated funding is assumed to create the conditions for this scientific elite to produce the kind of groundbreaking research that is desired in the competitive knowledge economy.

Centre of Excellence schemes form an example of such intended cumulative advantage. However, in an empirical study of Centers of Excellence in the Nordic countries, Langfeldt et al. (2015) found limited advantages from the Centre of Excellence status and funding for groups that were already high-performing. Somewhat lower-performing groups seemed to benefit more from the Centre of Excellence status as these showed a larger increase in publications and citations. Based on these findings, Langfeldt et al. (2015) conclude that there is a point of saturation to cumulative advantage.

Previous success in competitive funding programs is assumed to strengthen the position of these 'excellent' researchers in the acquisition of more funding. In this regard, Bol et al. (2018) found a Matthew effect when following applicants for an early career grant for 8 years after applying for funding: 'Over the entire period the lowest-ranked winners in the early career competition accumulated more than twice as much grant money as the highest-ranked nonwinners' (2018:4889). This can be partly ascribed to the use of past achievements as criteria in the evaluation of applications. But the authors also suggest a participation mechanism, whereby early career funding success encourages researchers to participate in future competitions.

Unintended concentration of funding may also result from funding organisations using relatively similar excellence criteria, while making grant decisions in isolation from one another (Aagaard et al. 2020). While the excellence regime prescribes that funding is allocated to the best of the best based on merit or quality, many factors outside the control of the researcher, and unrelated to the scientific quality of the work, have an effect on the success of grant applications. Laudel (2006) speaks in this regard of the 'quality myth' and hence argues against the use of 'research income' as quality indicator in evaluation procedures both at the level of institutions and of individual researchers. As Laudel (2006) suggests that a total research income is more likely to be the result of a Matthew effect than that it reflects the actual quality of the work.

Based on an extensive literature review on the debate around the relative concentration or dispersal of research funding, Aagaard et al. (2020) point at the limited evidence for the assumed advantages of concentration of resources in the current research ecosystem. They argue that most systems have taken the concentration of resources too far and that a greater dispersal of funding instead would allow a wider variety of researchers to conduct a wider variety of research, with more opportunities for original path-breaking ideas to emerge. In other words, the concentration of resources not only enhances existing inequalities through the Matthew effect, it is also understood to be rather counter-productive to scientific development (Bloch & Sörensen 2015).

3.2 Hyper-competition in the knowledge economy

Policies promoting an excellence regime assume that competition leads to better quality research. However, it has been argued that the emphasis on concentration of resources has led to hyper-competition and that this does not necessarily maximize quality. This body of critical literature presents excellence initiatives first and foremost as instruments of the wider neo-liberalization of academia, with hyper-competition as its perverse effect. Hyper-competition encourages a variety of negative behaviors among scientists including goal displacement, fraud, impact factor chasing and gaming strategies (e.g. Moore et al. 2017). Competition used as a

management tool, transforms the issue of quality into a competition for best rankings (Vessuri et al. 2013). Watermeyer and Olssen characterize the UK REF in this regard as a 'competitive market game' (2016: 203). Uses of notions of excellence in university's publicity materials and public speeches are often not substantiated by evidence of the acclaimed quality, and therefore better understood as advertorial strategies to mobilize competitive advantage and resources in a highly competitive environment (Wangenge-Ouma 2010).

Competition mechanisms, like rankings, produce exclusions. Differentiating between the best and the rest, can condemn the rest to being 'forgotten and neglected', whether these are individual researchers or countries (Vessuri et al. 2013). In the context of the UK 2014 REF, Watermeyer and Olssen (2016), discuss institutional strategic decision making in order to avoid harm to the institution's competitiveness, that excluded researchers and research from even participating in the REF.

Such gaming strategies directly impact the academic work environment: 'the consequence is an omnipresence of competitiveness that engenders repulsion, division, discomfort and fear, far more than it might incentivize, harmonize and instill a sense of belonging' (Watermeyer & Olssen 2016: 202). In a similar vein, Kamerlin (2015) warns that an overemphasis on competition damages collaborative work environments and installs a 'what's in it for me?' mentality in young researchers rather than a taste for collegiality. From this perspective, by fostering hyper-competition, the excellence regime privileges individual success at the cost of more collaborative approaches to science and knowledge production.

Rewarding the winners in a competitive funding environment is assumed to incite a trickle-down effect: all participants benefit as the quality of the work of the selective few improves the performance of the system as a whole. Scholten et al. (2021) indeed suggest that the losers are as much affected as the winners of competitions, albeit in a different way than intended by such policies. They observed that while only a small group benefits, and gains recognition and autonomy, all researchers experience the pressure to compete in the excellence regime and make decisions about priorities in anticipation of excellence funding.

This speaks to concerns about the excellence regime discouraging 'non-glamourous' work like replication studies (Moore et al. 2017) and about the tendency of evaluation systems to constrain critical and creative abilities in academic work (Watermeyer & Olssen 2016). Collectively these authors argue that the anticipated positive effects of competition are undermined by the way that the incentives generated by excellence funding arrangements negatively affect research cultures and decision making practices.

3.3 Homogeneity, conservatism and the reproduction of inequalities

Excellence initiatives and discourses are underpinned by meritocratic ideals. Competitive schemes dedicated to rewarding the best of the best operate under the presumption that opportunities for successful performance are equal and that excellence criteria can hence be considered 'objective' measures. Several authors have argued that in practice the excellence regime rather reproduces existing power structures: 'The faith in meritocracy is at the heart of how inequalities are reproduced' (Scully 2002 in Van den Brink & Benschop 2011). The excellence regime has been associated with a bias towards sameness which has consequences for the kinds of research and researchers that are more and less likely to get funded under excellence criteria.

In the literature addressing how the excellence regime fosters conservatism, the emphasis is on the reproduction of gender imbalances and biases and the favorable conditions for 'conventional' disciplinary research. But inequalities are also (re-)produced in other respects. Neylon for example points out that the implicit biases in notions of excellence mark differences in opportunities in geographic and racial terms: "The core problem of the rhetoric of research excellence is its homogeneity and its consequent privileging of North Atlantic and therefore inevitably white voices" (2020:105). The international predominance of excellence standards developed in Western European and North American contexts, are however more often discussed in terms of the local suitability of such standards. Again, it is in the presentation of excellence measures and criteria as 'neutral,' 'objective' and 'international' that the space is crafted for the reproduction and reinforcement of inequities in the research ecosystem.

For example, the presumed neutrality of excellence criteria has been argued to obscure gender-biased practices in the evaluation of candidates for faculty appointments. Vague notions of 'excellence' facilitate its use as a 'rationalizing myth' for selecting candidates based on gendered cognitive bias, personal relationships and ideas of 'local fit' that reproduce rather than challenge gender imbalances among faculty (O'Connor et al. 2020). It is precisely through the presentation of evaluation criteria, such as having an elaborate academic track record, as gender neutral that inequalities are reproduced (Van den Brink & Benschop 2011). Ideas of merit are strongly rooted in a history where women and other minoritized groups were excluded from academia. As discussed in section 1.1, it was eugenicist concerns about the decline in 'eminent men' that informed the production of academic performance measures in the early 20th century. Traces of this history can for example be observed in the types of activities and work/life balance that academics are expected to engage in (Thornton 2013; Kraemer-Mbula 2020). Present day criteria that constitute 'the ideal academic' are informed by what Thornton (2013) refers to as

'benchmark masculinity.' Importantly, gender is not the only marker of difference around which inequalities are reproduced in the excellence regime. The 'myth of colour-blind meritocracy' disadvantages racialized and Indigenous academics and leaves space for the perpetuation of racism precisely through considering racialized differences irrelevant in selection and assessment practices (Ahmed, 2012; Mohamad & Beagan 2019).

At the level of academic disciplines, excellence initiatives that come with a 'one model fits all' approach failing to take epistemic differences between disciplines into account, negatively affect the humanities in particular (Borlaug & Langfeldt 2020). Cañibano et al. (2018) note that, as an effect of the national evaluation system, researchers in history departments in Spain feel pushed to publish articles in high impact factor journals, while the discipline values monographs and books. This imposes a work time frame at odds with the temporality of the type of work historians engage in and is perceived as undermining disciplinary values of rigor and innovation.

Excellence measures have also been demonstrated to come with a bias against more risky and interdisciplinary research. Rafols et al. (2012) showed that the rather interdisciplinary field of Innovation Studies is at a relative disadvantage when excellence measures are based on journal rankings. Higher ranked journals are mono-disciplinary in nature and hence less accessible to Innovation Studies scholars for publication. Also in the context of funding allocation through peer review panels, lower success rates have been observed for interdisciplinary applications as compared to more 'conventional' applications in a scheme where these proposals were competing against each other (Langfeldt 2006).

Excellence measures have undervalued and excluded epistemic practices that do not fit the model of the research intensive university focused on the production of academic publications. This becomes particularly apparent in national contexts where attempts are made to accommodate Indigenous modes of knowing as part of reconciliation efforts. Research funding organisations in Canada, Australia and New Zealand for example have revised excellence and eligibility criteria and created direct programs and review practices to create space for Indigenous research practices in their funding portfolios.[16] Shifts include a stronger focus on collaboration, the research process and knowledge dissemination activities beyond academia. Such developments shed light on the narrow and exclusionary approach to knowledge production that the excellence regime tends to perpetuate while also showing potential for how the research ecosystem could be designed otherwise.

[16] See for example the 'Reference Group for the Appropriate Review of Indigenous Research' (https://cihr-irsc.gc.ca/e/52136.html), the Health Research Council of New Zealand's assessment criterion captured in the 'Maori Health Advancement Guidelines' (https://www.hrc.govt.nz/resources/maori-health-advancement-guidelines) and the inclusion of Indigenous research as a discipline in itself in the next Excellence in Research for Australia evaluation exercise (https://www.arc.gov.au/news-publications/media/feature-articles/release-era-ei-review-outcomes).

3.4 Between 'international' excellence and 'local' relevance

In the global knowledge economy, excellence is often considered in relation to a particular 'international standard.' In practice, this standard may take the shape of a metric (e.g. top percentile of most-cited papers) or expert judgment (i.e. peers considered knowledgeable about the global state of the art in a specific field). The extent to which 'international' excellence measures are suitable for assessment of research across cultural and geographical settings has been questioned by several authors. For example, output focused bibliometrics are often based on databases that lack coverage of publications in languages other than English. This considerably affects what research gets promoted and evaluated in the context of national evaluation systems (Lopez Piñeiro & Hicks 2015).

International excellence and local relevance tend to be presented as trade-offs and hence pose a challenge for funding systems outside the European and North-American centers of origin of dominant excellence measures. Framing excellence and relevance as 'contradictory', Radosevic and Lepori point out that funding systems 'have to endure world scientific excellence but should also be oriented towards needs of local stakeholders which are not necessarily contributing to world science but to local technology generation' (2009:664). The authors suggest that, for peripheral countries, achieving a balance between these different objectives would mean to find a suitable mix between promoting currently excellent fields and letting new areas of research emerge.

Vessuri et al. (2013) note that, in Latin-American countries, 'excellence' translates into doing well in competition with the North (see also Lorenz-Meyer 2012 about the Eastern European context) and that this regime may do more harm than good for enhancing the overall quality of science in the region. The authors highlight Latin-American initiatives that use the tools of international excellence measures for local purposes, using local libraries and calculations of citation scores. These interventions shift the focus from excellence in relation to a global system to improving science in and for Latin America. Anderson et al. (2015) similarly argue against the uncritical adoption of internationalized frameworks of excellence. The authors propose a conceptualization of science in terms of a 'knowledge dialogue model' that, while being in conversation with the global scientific community, accounts for local socio-ecological relevance and argue for the inclusion of 'broader impact' criteria in science policy and evaluation to incentivize researchers to focus on local challenges.

Funders with limited resources may be inclined to fund low-risk projects with more reliable returns on investment. Such a funding strategy fosters 'incremental science' rather than the cutting-edge, frontier projects that are associated with achieving 'world-class excellence' (Tijssen

& Kraemer Mbula 2018). In interviews with research coordinators of African Science Granting Councils, Tijssen and Kraemer-Mbula noticed in addition that interviewees often,

> *'highlighted the different views that international donors and national funding agencies have in terms of the performance parameters and indicators that are relevant and applicable to measure research quality and excellence. In this respect it was often noted that some of the indicators expected from international funding agencies are often non-existent or non-applicable in an African context'* (2018:397).

Next to breaking down the obstacles that currently prevent researchers from access to international publication, the authors argue for an expansion of excellence indicators beyond international publications to include a more diverse range of outputs that are relevant for the African context (cf. Chataway et al. 2019) and that take a longer term perspective on impact (Tijssen 2020).

Although international standards of excellence and local relevance of research are considered in tension with one another at the level of science policy and funding activities, Barnard et al. (2012) argue that there is no such trade-off in knowledge production at the level of individual scientists. Studying the internal and external connectedness of 'world-leading' South African scientists in the natural sciences, they observed that leading researchers in developing economies play an important role in the distribution of new knowledge in their respective countries and promote local science by employing both international and local researchers in collaborative projects. This suggests that rather than accepting the distinction and suggested tension between 'excellence' and 'relevance' at face-value, it is important to provide situated analyses of the dynamics of knowledge production practices.

4. Moving beyond excellence?

Discomfort around notions of excellence seem almost as ubiquitous as the concept itself. In interviews about the development of the European Commission's Research Excellence in Science and Technology indicator, Ferretti and colleagues (2018) found that policy makers, indicator developers and scholars were keen to point out the limitations of the use of the concept of excellence, but struggled to come up with alternatives. Similarly, alternatives to excellence models are less prevalent in the literature than critiques. In many cases, therefore, excellence continues to be at the center of the research ecosystem in general and the work of funders in particular. Some developments aim at improving or refining excellence initiatives and indicators, such as the scoreboard developed by Tijssen (2003), might be undertaken but these tend to leave the fundamentals of excellence unchanged.

However, several suggestions have been made to counter negative effects associated with the excellence regime. We have grouped these into three main categories: patching, pluralizing and transforming responses to excellence. The first aims to address obvious problems in the existing system associated with excellence, the second aims to widen the scope of excellence, the third aims to make fundamental changes to the research ecosystem in relation to excellence. These three categories are not mutually exclusive or exhaustive but rather are meant as sensitizing concepts illustrating differences in the way that mitigating strategies can be transformative of the research ecosystem.

4.1 Patching existing frameworks

Some influential changes to the excellence regime have been established through DORA (2012) and the Leiden Manifesto (Hicks et al. 2015).[17] These statements call for the abandonment of journal impact factors in favor of more responsible uses of metrics and have been widely taken up by research funding organisations and evaluation bodies to mitigate the overreliance on bibliometrics in decision making. While quantified approaches have been critiqued for narrowing down understandings of excellence, other authors suggest that more specific and refined bibliometrics hold the solution to time and cost intensive assessments based on peer review and could in some cases completely substitute such laborious procedures (e.g. Rodriquez Navarro & Brito 2020).

[17] Building on these statements, the more recently published Hong Kong principles (Moher et al. 2020) call in addition for evaluation practices in career assessment that focus on strengthening research integrity in order to counter questionable research practices induced by a regime of hyper-competition.

Other suggested patches aim to mitigate the negative side effects of excellence funding on individual researchers and research groups. For example, adjusting the size of the grant to the needs of different fields instead of the dominant one-model-fits-all approach (Borlaug & Langfeldt 2019) or making alternative funding available next to excellence schemes to mitigate strategic anticipation (Scholten et al. 2021). Such alterations do not so much challenge the meritocratic and competitive principles of the excellence regime but function as patches to emerging problems around particular performance measures and funding programs.

4.2 Pluralizing excellence

Several authors identify or suggest ways to diversify or pluralize the notion of excellence as a way to counter narrow understandings of it. Pluralization efforts at the level of academic outputs encompass, for example, the development of measures of openness and use of research to move bibliometric assessment away from traditional indicators such as journal impact factors.

Another strand of pluralization aims to move beyond the emphasis on published outputs, for example by making measures of impact part of excellence initiatives. The inclusion of 'impact case studies' to assess economic and societal impact of research in the UK REF and the controversies around it has received ample attention (e.g. Martin 2011; Smith et al. 2011; Penfield et al. 2014; Khazragui & Hudson 2015; Samuel & Derrick 2015; Williams & Grant 2018; Bandola-Gill 2019; Pinar & Unlu 2020). In the REF, impact is defined as 'an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia'.[18] Bandola-Gill (2019) argues that broadening the REF to include impact has highlighted the tensions between (academic) excellence and (policy) relevance and enhanced the division between categories of epistemic practices such as producing academic research, translating research, and producing policy research.

Other calls to pluralize the notion of excellence include, but are not limited to, suggestions that expand the excellence umbrella to include a measure of 'academic freedom' in both teaching and doing research as an indicator of excellence (Karran & Mallinson 2019), a parallel set of evaluation criteria for transdisciplinary research (Belcher et al. 2016) and the inclusion of impact measures specific to local contexts (e.g. Anderson et al. 2015; Tijssen & Kraemer-Mbula 2018; Chataway 2019; Tijssen 2020). The notion of 'inclusive excellence' gained track in North-American research and higher education institutions after 2005 when it was coined in a report commissioned by the Association of American Colleges and Universities to link quality and equity, diversity and inclusion (EDI) (Williams et al. 2005).

[18] https://re.ukri.org/research/ref-impact/.

Attention to EDI, indigenous rights, research cultures and academic work environments is also growing in the context of research funding activities. For example, the 'Tri-agency Statement on EDI' from the three federal research funding organisations in Canada promotes EDI as a prerequisite for achieving excellence: 'Achieving a more equitable, diverse and inclusive Canadian research enterprise is essential to creating the excellent, innovative and impactful research necessary to advance knowledge and understanding, and to respond to local, national and global challenges.'[19]

Such efforts to broaden notions of excellence are not (yet) reflected in the academic literature on research funding included in this review.[20] This emphasizes the relevance of including 'grey literature' such as policy papers, institutional reports and blog posts in considering current concerns around 'excellence' and also provides impetus to the empirical work linked to this paper in which we have undertaken analysis of the role of excellence in funding organisations through a set of detailed case studies, and which will be reported separately.

Whether pluralizing initiatives transform narrow definitions of excellence partly depends on whether impact, relevance and matters of EDI and Indigenous rights are proposed as separate measures next to excellence (leaving the narrow definition untouched) or included as indicators of excellence (broadening its meaning). However, the extent to which notions of excellence can be stretched and transform the excellence regime remains a matter of empirical investigation. The pluralization or differentiation of notions of excellence can help restore balance between the different tasks of researchers and institutions but also raises the concern of additional pressure put on researchers to fulfill criteria in different domains when these additional measures become subject to ranking and competition (Scholten et al. 2018).

4.3 Transforming the research ecosystem

Authors that hint at the need for radical transformation of the research ecosystem, sometimes offer broad conceptual suggestions for change. For example, it has been suggested that the singular focus on performance should be replaced by an understanding of science as an 'open, extended and complex system with a range of competing (and legitimate) perceptions of performance' (Rafols et al. 2012) or shifting from an international 'brain circulation' to a 'knowledge dialogue' model that also takes the local into account (Andersen et al. 2015). Others firmly call for the total abandonment of assessment criteria of 'productivity' (Halfman & Radder

[19] https://www.nserc-crsng.gc.ca/NSERC-CRSNG/EDI-EDI/index_eng.asp.
[20] Academic literature mainly discusses 'inclusive excellence' in the context of higher education (e.g. Sweeney 2013; Posselt 2014; Bleich et al. 2015; Harris et al. 2015).

2015) or even the abandonment of any notion of 'excellence' (e.g. Neylon 2020) as these are considered to be incommensurable with 'true' academic ideals.

Thinking more concretely through how excellence could be otherwise, Young (2015) distinguishes different uses of the notion of excellence in terms of 'zero-sum excellence' and 'threshold excellence' in the context of grant allocation. The former configures excellence as a relative concept, meaning 'the best of the best,' and thus comes with a moving target dependent on the number of applicants and available funds. The latter configures excellence as a quality of something assessed by a stable performance target. Additional and diverse criteria could then be used to further narrow the amount of applications down.

While most funding and evaluation systems operate with a 'zero-sum' version of excellence, Tijssen (2020) argues that threshold excellence is more compatible with distributive justice arguments and would better fit the working situation of science granting systems in the Global South. Moore et al. (2017) similarly call for a radical move away from a competitive funding system that is based on the evaluation of output and suggest a distribution of resources amongst all applicants that meet minimum criteria, for example through lottery. Concerns with (partial) randomization are that researchers fear for stigma and funders for the reputation of their schemes as it contradicts traditional merit-based decision making (Bendiscoli 2019). Considering these reserves and that radical transformation takes time, Ioannidis (2011) suggests starting with pilot projects to experiment with a list of alternative funding strategies.

Notwithstanding these examples, in common with Ferriti et al.'s (2018) policy makers and indicator developers, authors are generally keen to criticize, while alternatives to the performance-based hyper-competitive funding and evaluation systems operating with a zero-sum configuration of excellence, are only scarcely articulated.

5. Discussion and conclusion

The notion of excellence has a history. While this history is often narrated as rather recent, starting with 'excellence' entering science policies in the 1990s, in the first section of this paper we traced excellence back to early 20th century eugenicist concerns around the decline in eminent men and the first university rankings. We have also placed the notion of scientific excellence in the wider context of excellence in management and business, which developed from the 1980s onwards. However, a thorough historical study could shed much greater light on the emergence, continuity and discontinuity of notions of excellence.

One relationship to further explore would be that between scientific excellence and the notions of excellence in the field of business management. This could provide insight into how the overlap between these discourses in the context of policy making has shaped notions of excellence in the research ecosystem. The history of excellence also locates it in space. The North-American and Western European origins of its emergence and dominance in the research ecosystem also become apparent in the geographical distribution of research on excellence initiatives. Although this literature review includes references to studies conducted in a range of settings across the globe, the European context is overrepresented (although this is reinforced by our limitation to English language literature).

Excellence remains an ill-defined concept. The vagueness surrounding notions of excellence has performative effects. It can be used as a 'rationalising myth' obscuring other motivations (O'Connor et al. 2020) and as a 'quintessential polymorphic term' excellence gains different meanings in different situations (Lamont 2009:159). This also allows excellence to mean and do something else in relation to evaluation, funding and research practices respectively. For example, how the excellence regime affects opportunities for interdisciplinary research has been found to be positive (Centers of Excellence, Hellström et al. 2018) as well as negative (journal publications, Rafols 2012). The ERC excellence funding scheme has been associated with risk taking (Laudel & Gläser 2014) as well as with risk aversion (Luukkonen 2012).

The excellence regime in general has been considered to discourage confirmatory research such as replication studies (Moore et al. 2017). These apparently somewhat contradicting findings indicate that constitutive effects in the excellence regime look rather different depending on where and how these are studied. In addition, the omnipresence of notions of excellence and the ambiguities surrounding them, complicates the general attribution of effects in the excellence regime. Rather than coming to crude generalizations or trying to grasp what excellence essentially is, situated empirical analyses of notions of excellence that address what 'excellence'

becomes and does in different practices would be particularly helpful in moving the discussion forward.

The closer the focus on the everyday activities of researchers and research groups, the more apparent the constitutive effects of a competition-driven excellence regime become. Several authors have argued that there is an imbalance between intended and unintended consequences of competition and concentration of resources on what are supposedly 'the best of the best' (e.g. Aagard 2020; Scholten et al. 2021).

However, we prefer to think in terms of 'constitutive effects' (Dahler-Larsson 2014) rather than aim to make a distinction between intended and unintended consequences as it is not so easy to disentangle what effects are intended or unintended and by whom, or what precisely these are a consequence of. The concept of constitutive effects helps shift attention to *how* mobilizations of the notion of excellence have shaped the research ecosystem and how these particular mobilizations (re)produce inequalities. For instance, how a strong focus on academic publications in the form of journal articles has shaped epistemic practices to the exclusion of other forms of knowledge production and dissemination.

While purely merit-based allocation of funding has been shown to be something of an illusion, alternatives are only slowly emerging and hardly challenging Mertonian understandings of excellence and recognition deeply ingrained in the current research ecosystem. Rather than moving away from competition-based systems, the pluralization of notions of excellence has recently received wider attention. Such pluralizing for example includes EDI goals as performance measures and increased attention for the 'impact' of research beyond academia.

This raises the question to what extent notions of excellence can be stretched. What sticks to excellence? Will a broadening of indicators and criteria be sufficient to mitigate undesirable consequences or will the problems associated with competition and meritocratic ideals that underpin notions of excellence persist or themselves be broadened? While these assumptions seem to remain largely unquestioned in the literature, ideals around competition and meritocracy are foundational to mobilizations of excellence in the research ecosystem and hence require as much critical attention.

Different notions of excellence co-exist in the research ecosystem (cf. Langfeldt et al. 2019). Research funding organisations play an important role in the institutionalization of excellence discourses through the development and execution of competitive funding mechanisms and evaluation regimes. It is therefore surprising that in the empirical research on the excellence regime the activities of research funding organisations themselves have been underexplored. Langfeldt et al. (2019) note a differentiation between those funding organisations that aim to set the 'gold standard' for research quality and those that tend to follow, but do not explicate how

this differentiation precisely comes about or affects the research ecosystem more broadly. Research funding organisations play an important role in the proliferation of existing excellence discourses and execution of science policies, reflecting the values of their communities.

Moreover, research funding organisations are also actively defining, developing and transforming standards and criteria of research excellence and accordingly intervene in the research ecosystem. As sites of intervention in which policy, funding and research activities come together, research funding organisations make an interesting case to further explore the future of 'excellence.' Of course, these organisations cannot do this alone, and so need to engage in a wide-ranging dialogue with other actors in the research ecosystem in order to work out solutions that will meet with wide acceptance and will strengthen the research endeavor.

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: 106–17. DOI: 10.1093/reseval/rvv003

Aagaard, Kaare, Kladakis, A., & Nielsen, M. W. (2020). 'Concentration or dispersal of research funding?', *Quantitative Science Studies*, 1/1: 117–49. DOI: 10.1162/qss_a_00002

Abramo, G., D'Angelo, C. A., & Costa, F. D. (2011). 'National research assessment exercises: a comparison of peer review and bibliometrics rankings', *Scientometrics*, 89/3: 929–41. DOI: 10.1007/s11192-011-0459-x

Ackers, L. (2008). 'Internationalisation, Mobility and Metrics: A New Form of Indirect Discrimination?', *Minerva: A Review of Science, Learning & Policy*, 46/4: 411–35. DOI: 10.1007/s11024-008-9110-2

Ahmed, S. (2012). *On being included*. Durham and London: Duke University Press. DOI: 10.1515/9780822395324

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). 'Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories', *SAGE Open*, 9/1: 2158244019829575. DOI: 10.1177/2158244019829575

Anderson, C. B., Monjeau, A., & Rau, J. R. (2015). 'Knowledge Dialogue to Attain Global Scientific Excellence and Broader Social Relevance', *BioScience*, 65/7: 709–17. DOI: 10.1093/biosci/biv046

Antonowicz, D., Kohoutek, J., Pinheiro, R., & Hladchenko, M. (2017). 'The roads of "excellence" in Central and Eastern Europe', *European Educational Research Journal*, 16/5: 547–67. DOI: 10.1177/1474904116683186

Araújo, L. S., Cruz, J. F. A., & Almeida, L. S. (2017). 'Achieving scientific excellence: An exploratory study of the role of emotional and motivational factors', *High Ability Studies*, 28/2: 249–64. DOI: 10.1080/13598139.2016.1264293

van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2014). 'The selection of talent as a group process. A literature review on the social dynamics of decision making in grant panels', *Research Evaluation*, 23/4: 298–311. DOI: 10.1093/reseval/rvu017

Auranen, O., & Nieminen, M. (2010). 'University research funding and publication performance—An international comparison', *Research Policy*, 39/6: 822–34. DOI: 10.1016/j.respol.2010.03.003

Balbuena, L. D. (2018). 'The UK Research Excellence Framework and the Matthew effect: Insights from machine learning', *PLOS ONE*, 13/11: e0207919. DOI: 10.1371/journal.pone.0207919

Bandola-Gill, J. (2019). 'Between relevance and excellence? Research impact agenda and the production of policy knowledge', *Science and Public Policy*, 46/6: 895–905. DOI: 10.1093/scipol/scz037

Barnard, H., Cowan, R., & Müller, M. (2012). 'Global excellence at the expense of local diffusion, or a bridge between two worlds? Research in science and technology in the developing world', *Research Policy*, 41/4: 756–69. DOI: 10.1016/j.respol.2011.12.002

Belcher, B. M., Rasmussen, K. E., Kemshaw, M. R., & Zornes, D. A. (2016). 'Defining and assessing research quality in a transdisciplinary context', *Research Evaluation*, 25/1: 1–17. DOI: 10.1093/reseval/rvv025

Bendiscioli, S. (2019). 'The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision-making', *EMBO reports*, 20/12. DOI: 10.15252/embr.201949472

van den Besselaar, P., & Sandström, U. (2015). 'Early career grants, performance, and careers: A study on predictive validity of grant decisions', *Journal of Informetrics*, 9/4: 826–38. DOI: 10.1016/j.joi.2015.07.011

van den Besselaar, P., Sandström, U., & Schiffbaenker, H. (2018). 'Studying grant decision-making: a linguistic analysis of review reports', *Scientometrics*, 117/1: 313–29. DOI: 10.1007/s11192-018-2848-x

Bleich, M. R., MacWilliams, B. R., & Schmidt, B. J. (2015). 'Advancing Diversity Through Inclusive Excellence in Nursing Education', *Journal of Professional Nursing*, 31/2: 89–94. DOI: 10.1016/j.profnurs.2014.09.003

Bloch, C., & Sorensen, M. P. (2015). 'The size of research funding: Trends and implications', *Science and Public Policy*, 42/1: 30–43. DOI: 10.1093/scipol/scu019

Bol, T., Vaan, M. de, & Rijt, A. van de. (2018). 'The Matthew effect in science funding', *Proceedings of the National Academy of Sciences*, 115/19: 4887–90. DOI: 10.1073/pnas.1719557115

Bonaccorsi, A., Cicero, T., Haddawy, P., & Hassan, S.-U. (2017). 'Explaining the transatlantic gap in research excellence', *Scientometrics*, 110/1: 217–41. DOI: 10.1007/s11192-016-2180-2

Borlaug, S. B. (2016). 'Moral hazard and adverse selection in research funding: Centres of excellence in Norway and Sweden', *Science and Public Policy*, 43/3: 352–62. DOI: 10.1093/scipol/scv048

Borlaug, S. B., & Langfeldt, L. (2020). 'One model fits all? How centres of excellence affect research organisation and practices in the humanities', *Studies in Higher Education*, 45/8: 1746–57. DOI: 10.1080/03075079.2019.1615044

Bougnol, M.-L., & Dulá, J. H. (2015). 'Technical pitfalls in university rankings', *Higher Education*, 69/5: 859–66. Springer.

Brankovic, J., Ringel, L., & Werron, T. (2018). 'How Rankings Produce Competition: The Case of Global University Rankings', *Zeitschrift für Soziologie*, 47/4: 270–88. De Gruyter. DOI: 10.1515/zfsoz-2018-0118

van den Brink, M., & Benschop, Y. (2012). 'Gender practices in the construction of academic excellence: Sheep with five legs', *organisation*, 19/4: 507–24. DOI: 10.1177/1350508411414293

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: 775–89. DOI: 10.1093/scipol/scy025

Chataway, J., Dobson, C., Daniels, C., Byrne, R., Hanlin, R., & Tigabu, A. (2019). 'Science granting councils in Sub-Saharan Africa: Trends and tensions', *Science and Public Policy*, 46/4: 620–31. DOI: 10.1093/scipol/scz007

Cremonini, L., Horlings, E., & Hessels, L. K. (2018). 'Different recipes for the same dish: Comparing policies for scientific excellence across different countries', *Science and Public Policy*, 45/2: 232–45. DOI: 10.1093/scipol/scx062

Cruz-Castro, L., Benitez-Amado, A., & Sanz-Menéndez, L. (2016). 'The proof of the pudding: University responses to the European Research Council', *Research Evaluation*, 25/4: 358–70. DOI: 10.1093/reseval/rvw015

Dahler-Larsen, P. (2014). 'Constitutive Effects of Performance Indicators: Getting beyond unintended consequences', *Public Management Review*, 16/7: 969–86. Routledge. DOI: 10.1080/14719037.2013.770058

Degn, L., Franssen, T., Sørensen, M. P., & de Rijcke, S. (2018). 'Research groups as communities of practice—a case study of four high-performing research groups', *Higher Education (00181560)*, 76/2: 231–46. DOI: 10.1007/s10734-017-0205-2

Ferretti, F., Pereira, Â. G., Vértesy, D., & Hardeman, S. (2018). 'Research excellence indicators: time to reimagine the "making of"?', *Science and Public Policy*, 45/5: 731–41. DOI: 10.1093/scipol/scy007

Fisher, D., Atkinson-Grosjean, J., & House, D. (2001). 'Changes in Academy/Industry/State Relations in Canada: The Creation and Development of the Networks of Centres of Excellence', *Minerva: A Review of Science, Learning & Policy*, 39/3: 299–325. DOI: 10.1023/A:1017924027522

Flink, T., & Peter, T. (2018). 'Excellence and Frontier Research as Travelling Concepts in Science Policymaking', *Minerva*, 56/4: 431–52. DOI: 10.1007/s11024-018-9351-7

Fochler, M., Felt, U., & Müller, R. (2016). 'Unsustainable Growth, Hyper-Competition, and Worth in Life Science Research: Narrowing Evaluative Repertoires in Doctoral and Postdoctoral Scientists' Work and Lives', *Minerva*, 54/2: 175–200. DOI: 10.1007/s11024-016-9292-y

Franssen, T., & De Rijcke, S. (2019). 'The rise of project funding and its effects on the social structure of academia'. Cannizzo F. & Osbaldiston N. (eds) *The social structures of global academia*, pp. 144–61. Routledge: London.

Franssen, T., Scholten, W., Hessels, L. K., & de Rijcke, S. (2018). 'The Drawbacks of Project Funding for Epistemic Innovation: Comparing Institutional Affordances and Constraints of Different Types of Research Funding', *Minerva*, 56/1: 11–33. DOI: 10.1007/s11024-017-9338-9

Franssen, T., & Wouters, P. (2019). 'Science and its significant other: Representing the humanities in bibliometric scholarship', *Journal of the Association for Information Science and Technology*, 70/10: 1124–37. DOI: https://doi.org/10.1002/asi.24206

Fu, Y.-C., Baker, D. P., & Zhang, L. (2020). 'Engineering a World Class University? The Impact of Taiwan's World Class University Project on Scientific Productivity', *Higher Education Policy*, 33/3: 555–70. DOI: 10.1057/s41307-018-0110-z

Garfield, E. (2007). 'The evolution of the Science Citation Index', *International Microbiology*, 10: 65–9. DOI: 10.2436/20.1501.01.10

Gläser, J., & Velarde, K. S. (2018). 'Changing Funding Arrangements and the Production of Scientific Knowledge: Introduction to the Special Issue', *Minerva*, 56/1: 1–10. DOI: 10.1007/s11024-018-9344-6

Godin, B. (2007). 'From Eugenics to Scientometrics: Galton, Cattell, and Men of Science', *Social Studies of Science*, 37/5: 691–728. SAGE Publications Ltd. DOI: 10.1177/0306312706075338

Halffman, W., & Radder, H. (2015). 'The Academic Manifesto: From an Occupied to a Public University', *Minerva: A Review of Science, Learning & Policy*, 53/2: 165–87. DOI: 10.1007/s11024-015-9270-9

Hammarfelt, B., Rijcke, S., & Wouters, P. (2017). 'From Eminent Men to Excellent Universities: University Rankings as Calculative Devices', *Minerva: A Review of Science, Learning & Policy*, 55/4: 391–411. DOI: 10.1007/s11024-017-9329-x

Harris, J. C., Barone, R. P., & Davis, L. P. (2015). 'Who Benefits?: A Critical Race Analysis of the (D)Evolving Language of Inclusion in Higher Education', *Thought & Action*. National Education Association.

Hellström, T. (2011). 'Homing in on excellence: Dimensions of appraisal in Center of Excellence program evaluations', *Evaluation*, 17/2: 117–31. DOI: 10.1177/1356389011400891

——. (2018). 'Centres of Excellence and Capacity Building: from Strategy to Impact', *Science and Public Policy*, 45/4: 543–52. DOI: 10.1093/scipol/scx082

Hellström, T., Brattström, E., & Jabrane, L. (2018). 'Governing interdisciplinary cooperation in Centers of Excellence', *Studies in Higher Education*, 43/10: 1763–77. DOI: 10.1080/03075079.2018.1520476

Hellström, T., Jabrane, L., & Brattström, E. (2018). 'Center of excellence funding: Connecting organisational capacities and epistemic effects', *Research Evaluation*, 27/2: 73–81. DOI: 10.1093/reseval/rvx043

Hicks, D. (2012). 'Performance-based university research funding systems', *Research Policy*, 41/2: 251–61. DOI: 10.1016/j.respol.2011.09.007

Hicks, D., & Katz, J. S. (2011). 'Equity and Excellence in Research Funding', *Minerva*, 49/2: 137–51. DOI: 10.1007/s11024-011-9170-6

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). 'Bibliometrics: The Leiden Manifesto for research metrics', *Nature News*, 520/7548: 429. DOI: 10.1038/520429a

Hoenig, B. (2017). *Europe's New Scientific Elite: Social Mechanisms of Science in the European Research Area*. London: Routledge. DOI: 10.4324/9781315446042

Hornbostel, S., Böhmer, S., Klingsporn, B., Neufeld, J., & Ins, M. von. (2009). 'Funding of young scientist and scientific excellence', *Scientometrics*, 79/1: 171–90. DOI: 10.1007/s11192-009-0411-5

Hou, A. Y.-C., Ince, M., & Chiang, C.-L. (2012). 'A reassessment of Asian pacific excellence programs in higher education: the Taiwan experience', *Scientometrics*, 92/1: 23–42. DOI: 10.1007/s11192-012-0727-4

Ioannidis, J. P. A. (2011). 'Fund people not projects', *Nature*, 477/7366: 529–31. Nature Publishing Group. DOI: 10.1038/477529a

Jang, D., Doh, S., Kang, G.-M., & Han, D.-S. (2017). 'Impact of Alumni Connections on Peer Review Ratings and Selection Success Rate in National Research', *Science, Technology, & Human Values*, 42/1: 116–43. SAGE Publications Inc. DOI: 10.1177/0162243916665466

Johnes, J. (2018). 'University rankings: What do they really show?', *Scientometrics*, 115/1: 585–606. Springer.

Kamerlin, S. C. L. (2015). 'Hypercompetition in biomedical research evaluation and its impact on young scientist careers', *International Microbiology: The Official Journal of the Spanish Society for Microbiology*, 18/4: 253–61. DOI: 10.2436/20.1501.01.257

Karran, T., & Mallinson, L. (2019). 'Academic Freedom and World-Class Universities: A Virtuous Circle?', *Higher Education Policy*, 32/3: 397–417. DOI: 10.1057/s41307-018-0087-7

Kehm, B. (2013). 'To be or not to be? The impacts of the excellence initiative on the German system of higher education.' Shin J. C. & Kehm B. (eds) *Institutionalization of world-class university in global competition.*, pp. 81–97. Springer: Dordrecht.

Khazragui, H., & Hudson, J. (2015). 'Measuring the benefits of university research: impact and the REF in the UK', *Research Evaluation*, 24/1: 51–62. DOI: 10.1093/reseval/rvu028

Kraemer-Mbula, E. (2020). 'Gender diversity and the transformation of research excellence'. Kraemer-Mbula E., Tijssen R., Wallace M. L., & McClean R. (eds) *Transforming Research Excellence: New Ideas from the Global South*, pp. 79–91. African Minds: Capetown, South Africa.

König, T. (2017). *The European Research Council*. Cambridge, MA: Polity Press.

Lamont, M. (2009). *How Professor Think: Inside the Curious World of Academic Judgment.* Cambridge, MA: Harvard University Press.

Langfeldt, L. (2006). 'The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments', *Research Evaluation*, 15/1: 31–41. DOI: 10.3152/147154406781776039

Langfeldt, L., Benner, M., Sivertsen, G., Kristiansen, E. H., Aksnes, D. W., Borlaug, S. B., Hansen, H. F., et al. (2015). 'Excellence and growth dynamics: A comparative study of the Matthew effect', *Science and Public Policy*, 42/5: 661–75. DOI: 10.1093/scipol/scu083

Langfeldt, L., Nedeva, M., Sörlin, S., & Thomas, D. A. (2020). 'Co-existing Notions of Research Quality: A Framework to Study Context-specific Understandings of Good Research', *Minerva*, 58/1: 115–37. DOI: 10.1007/s11024-019-09385-2

Laudel, G. (2006). 'The "Quality Myth": Promoting and Hindering Conditions for Acquiring Research Funds', *Higher Education*, 52/3: 375–403. DOI: 10.1007/s10734-004-6414-5

Laudel, G., & Gläser, J. (2014). 'Beyond breakthrough research: Epistemic properties of research and their consequences for research funding', *Research Policy*, 43/7: 1204–16. DOI: 10.1016/j.respol.2014.02.006

López Piñeiro, C., & Hicks, D. (2015). 'Reception of Spanish sociology by domestic and foreign audiences differs and has consequences for evaluation', *Research Evaluation*, 24/1: 78–89. DOI: 10.1093/reseval/rvu030

Lorenz-Meyer, D. (2012). 'Locating Excellence and Enacting Locality', *Science, Technology, & Human Values*, 37/2: 241–63. DOI: 10.1177/0162243911409249

Lovakov, A., Panova, A., Sterligov, I., & Yudkevich, M. (2021). 'Does government support of a few leading universities have a broader impact on the higher education system? Evaluation of the Russian University Excellence Initiative', *Research Evaluation*, 00/0: 16.

Luukkonen, T. (2012). 'Conservatism and risk-taking in peer review: Emerging ERC practices', *Research Evaluation*, 21/1: 48–60. DOI: 10.1093/reseval/rvs001

——. (2014). 'The European Research Council and the European research funding landscape', *Science and Public Policy*, 41/1: 29–43. DOI: 10.1093/scipol/sct031

Martin, B. R. (2011). 'The Research Excellence Framework and the "impact agenda": are we creating a Frankenstein monster?', *Research Evaluation*, 20/3: 247–54. DOI: 10.3152/095820211X13118583635693

Mazzucato, M. (2018). *Mission-oriented research & innovation in the European Union*.

Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.

Mi Dahlgaard-Park, S., & Dahlgaard, J. J. (2007). 'Excellence – 25 years evolution', (S. Mi Dahlgaard-Park, Ed.)*Journal of Management History*, 13/4: 371–93. Emerald Group Publishing Limited. DOI: 10.1108/17511340710819606

Mohamed, T. & Beagan, B. L. (2019). ''Strange faces' in the academy: experiences of racialized and Indigenous faculty in Canadian universities', *Race Ethnicity and Education*, 22/3: 338-354. DOI: 10.1080/13613324.2018.1511532

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham M. H., Barbour V., et al. (2020) 'The Hong KongPrinciples for assessing researchers: Fostering research integrity', *PLoS Biology*, 18/7: e3000737.

DOI: 10.1371/journal.pbio.3000737Moore, S., Neylon, C., Paul Eve, M., Paul O'Donnell, D., & Pattinson, D. (2017). '"Excellence R Us": university research and the fetishisation of excellence', *Palgrave Communications*, 3/1. DOI: 10.1057/palcomms.2016.105

Münch, R. (2014). *Academic capitalism: Universities in the global struggle for excellence*. Routledge.

Neylon, C. (2020). 'Research excellence is a neo-colonial agenda (and what might be done about it)'. Kraemer-Mbula E., Tijssen R., Wallace M. L., & McLean R. (eds) *Transforming Research Excellence: New Ideas from the Global South*, pp. 92–115. African Minds: Capetown, South Africa.

O'Connor, P., López, E. M., Hagan, C. O., Wolffram, A., Aye, M., Chizzola, V., Mich, O., et al. (2020). 'Micro-political practices in higher education: a challenge to excellence as a rationalising myth?', *Critical Studies in Education*, 61/2: 195–211. DOI: 10.1080/17508487.2017.1381629

OECD. (2014). *Promoting Research Excellence: New Approaches to Funding*. OECD. DOI: 10.1787/9789264207462-en

Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). 'Assessment, evaluations, and definitions of research impact: A review', *Research Evaluation*, 23/1: 21–32. DOI: 10.1093/reseval/rvt021

Pinar, M. (2020). 'It is not all about performance: Importance of the funding formula in the allocation of performance-based research funding in England', *Research Evaluation*, 29/1: 100–19. DOI: 10.1093/reseval/rvz030

Pinar, M., & Unlu, E. (2020). 'Evaluating the potential effect of the increased importance of the impact component in the Research Excellence Framework of the UK', *British Educational Research Journal*, 46/1: 140–60. DOI: 10.1002/berj.3572

Posselt, J. R. (2014). 'Toward Inclusive Excellence in Graduate Education: Constructing Merit and Diversity in PhD Admissions', *American Journal of Education*, 120/4: 481–514. The University of Chicago Press. DOI: 10.1086/676910

Radosevic, S., & Lepori, B. (2009). 'Public research funding systems in central and eastern Europe: between excellence and relevance: introduction to special section', *Science and Public Policy*, 36/9: 659–66. DOI: 10.3152/030234209X475236

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). 'How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management', *Research Policy*, Exploring the Emerging Knowledge Base of 'The Knowledge Society', 41/7: 1262–82. DOI: 10.1016/j.respol.2012.03.015

Rijcke, S. de, 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: 161–9. DOI: 10.1093/reseval/rvv038

Rodríguez-Navarro, A., & Brito, R. (2020). 'Like-for-like bibliometric substitutes for peer review: Advantages and limits of indicators calculated from the ep index', *Research Evaluation*, 29/2: 215–30. DOI: 10.1093/reseval/rvaa002

Samuel, G. N., & Derrick, G. E. (2015). 'Societal impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014', *Research Evaluation*, 24/3: 229–41. DOI: 10.1093/reseval/rvv007

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: 364–78. DOI: 10.1057/eps.2009.19

Schneider, J. W., & van Leeuwen, T. N. (2014). 'Analysing robustness and uncertainty levels of bibliometric performance statistics supporting science policy. A case study evaluating Danish postdoctoral funding', *Research Evaluation*, 23/4: 285–97. DOI: 10.1093/reseval/rvu016

Scholten, W., Van Drooge, L., & Diederen, P. (2018). Excellence is extra-ordinary – Thirty years of focus on excellence in Dutch science policy. The Hague: Rathenau Instituut

Scholten, W., Franssen, T. P., van Drooge, L., de Rijcke, S., & Hessels, L. K. (2021). 'Funding for few, anticipation among all: Effects of excellence funding on academic research groups', *Science and Public Policy*, scab018. DOI: 10.1093/scipol/scab018

Schroder, S., Welter, F., Leisten, I., Richert, A., & Jeschke, S. (2014). 'Research performance and evaluation--Empirical results from collaborative research centers and clusters of excellence in Germany', *Research Evaluation*, 23/3: 221–32. DOI: 10.1093/reseval/rvu010

Sivertsen, G. (2016). 'Publication-Based Funding: The Norwegian Model'. Ochsner M., Hug S. E., & Daniel H.-D. (eds) *Research Assessment in the Humanities: Towards Criteria and Procedures*, pp. 79–90. Springer International Publishing: Cham. DOI: 10.1007/978-3-319-29016-4_7

——. (2017). 'Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective', *Palgrave Communications*, 3/1: 17078. DOI: 10.1057/palcomms.2017.78

Smith, S., Ward, V., & House, A. (2011). '"Impact" in the proposals for the UK's Research Excellence Framework: Shifting the boundaries of academic autonomy', *Research Policy*, 40/10: 1369–79. DOI: 10.1016/j.respol.2011.05.026

Sørensen, M. P., Bloch, C., & Young, M. (2016). 'Excellence in the knowledge-based economy: from scientific to research excellence', *European Journal of Higher Education*, 6/3: 217–36. DOI: 10.1080/21568235.2015.1015106

Star, S. L., & Griesemer, J. R. (1989). 'Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39', *Social Studies of Science*, 19/3: 387–420. SAGE Publications Ltd. DOI: 10.1177/030631289019003001

Staniland, N. A., Harris, C., & Pringle, J. K. (2020). ''Fit' for whom? Career strategies of Indigenous (Maori) Academics', *Higher Education,* 79: 589-604. DOI: 10.1007/s10734-019-00425-0

Thomas, D. A., Nedeva, M., Tirado, M. M., & Jacob, M. (2020). 'Changing research on research evaluation: A critical literature review to revisit the agenda', *Research Evaluation*, 29/3: 275–88. DOI: 10.1093/reseval/rvaa008

Thornton, M. (2013). 'The Mirage of Merit', *Australian Feminist Studies*, 28/76: 127–43. Routledge. DOI: 10.1080/08164649.2013.789584

Tijssen, R. (2020). 'Re-valuing research excellence: From excellentism to responsible assessment'. Kraemer-Mbula E., Tijssen R., Wallace M. L., & McClean R. (eds) *Transforming Research Excellence: New Ideas from the Global South*, pp. 59–78. African Minds: Capetown, South Africa.

Tijssen, R. J. (2003). 'Scoreboards of research excellence', *Research Evaluation*, 12/2: 91–103.

Tijssen, R., & Kraemer-Mbula, E. (2018). 'Research excellence in Africa: Policies, perceptions, and performance', *Science and Public Policy*, 45/3: 392–403. DOI: 10.1093/scipol/scx074

Vessuri, H., Guédon, J.-C., & Cetto, A. M. (2014). 'Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development', *Current Sociology*, 62/5: 647–65. DOI: 10.1177/0011392113512839

Wangenge-Ouma, G., & Langa, P. (2010). 'Universities and the mobilization of claims of excellence for competitive advantage', *Higher Education (00181560)*, 59/6: 749–64. DOI: 10.1007/s10734-009-9278-x

Watermeyer, R. (2016). 'Impact in the REF: issues and obstacles', *Studies in Higher Education*, 41/2: 199–214. Routledge. DOI: 10.1080/03075079.2014.915303

Watermeyer, R., & Olssen, M. (2016). '"Excellence" and Exclusion: The Individual Costs of Institutional Competitiveness', *Minerva: A Review of Science, Learning & Policy*, 54/2: 201–18. DOI: 10.1007/s11024-016-9298-5

Wigren-Kristoferson, C., Gabrielsson, J., & Kitagawa, F. (2011). 'Mind the gap and bridge the gap: Research excellence and diffusion of academic knowledge in Sweden', *Science and Public Policy*, 38/6: 481–92.

Williams, D. A., Berger, J. B., & McClendon, S. A. (n.d.). 'Toward a Model of Inclusive Excellence and Change in Postsecondary Institutions', 49.

Williams, K., & Grant, J. (2018). 'A comparative review of how the policy and procedures to assess research impact evolved in Australia and the UK', *Research Evaluation*, 27/2: 93–105. DOI: 10.1093/reseval/rvx042

Wilsdon, J., et al. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management.* DOI: 10.13140/RG.2.1.4929.1363

Wouters, P. (1999). *The Citation Culture* (Dissertation). University of Amsterdam, Amsterdam.

Yonezawa, A., & Shimmi, Y. (2016). 'Transformation of University Governance through Internationalization'. Liu N. C., Cheng Y., & Wang Q. (eds) *Matching Visibility and Performance: A Standing Challenge for World-Class Universities*, pp. 103–18. SensePublishers: Rotterdam. DOI: 10.1007/978-94-6300-773-3_6

Young, M. (2015). 'Shifting Policy Narratives in Horizon 2020', *Journal of Contemporary European Research*, 11/1.

Zhao, K., & You, Z. (2019). 'Isomorphism, Diversification, and Strategic Ambiguity: Goal Setting of Chinese Higher Education Institutions in the Double World-Class Project', *Higher Education Policy*. DOI: 10.1057/s41307-019-00168-8

Zong, X., & Zhang, W. (2019). 'Establishing world-class universities in China: deploying a quasi-experimental design to evaluate the net effects of Project 985', *Studies in Higher Education*, 44/3: 417–31. DOI: 10.1080/03075079.2017.1368475