

# Excellence R Us: University Research and the Fetishisation of Excellence

Samuel Moore, Cameron Neylon, Martin Paul Eve, Daniel Paul O'Donnell, Damian Pattinson<sup>1</sup>

## Abstract

The rhetoric of “excellence” is pervasive across the academy. It is used to refer to research outputs as well as researchers, theory and education, individuals and organisations, from art history to zoology. But what does “excellence” mean? Does it in fact mean anything at all? And is the pervasive narrative of excellence and competition a good thing? Drawing on a range of sources we interrogate “excellence” as a concept and find that it has no intrinsic meaning as used in the academy. Rather it functions as a linguistic interchange mechanism or boundary object. To investigate whether linguistic function is useful we examine how excellence rhetoric combines with narratives of scarcity and competition and show that hypercompetition that arises leads to a performance of “excellence” that is completely at odds with the qualities of good research. We trace the roots of issues in reproducibility, fraud, as well as diversity to the stories we tell ourselves as researchers and offer an alternative rhetoric based on soundness. “Excellence” is not excellent, it is a pernicious and dangerous rhetoric that undermines the very foundations of good research and scholarship.

## Introduction: The Ubiquity of Excellence Rhetoric

“Excellence” is the gold standard of the University world. Institutional mission statements proclaim, in almost identical language, their “international reputation for [educational] excellence” (Baylor, Imperial College London, Loughborough University, Monash University, The University of Sheffield are amongst many examples of institutions using these words), or the extent to which they are guided by principles of “excellence” (University of Cambridge, Carnegie Mellon, Gustav Adolphus, University College London, Warwick, again, among many). University

---

<sup>1</sup> In keeping with our argument, and following in an extensive tradition of subverting traditional scarce markers of prestige, the authors have adopted a redistributive approach to the order of their names in the byline. As an international collaboration of uniformly nice people (cf. Moran, Hoover, and Bestiale 2016; Hoover, Posch, and Bestiale 1987; Hoover et al. 1988; see Tartamella 2014 for an explanation), lacking access to a croquet field (cf. Hassell and May 1974), and not identifying any excellent pun to be made from ordering our names (cf. Alpher, Bethe, and Gamow 1948; Lord, de Vader, and Alliger 1986), we elected to assign index numbers based on alphabetical ordering by surname and to randomise these using an online tool. For the avoidance of doubt, while several of the authors have pets none of them are included as authors (cf. Matzinger and Mirkwood 1978). None of us are approaching a tenure decision (cf. Roderick and Gillespie 1998). And none of us are fictional entities who generate their papers algorithmically using SciGen (see Labbé 2010 for the contrasting case of Ike Antkare nevertheless greatly outranked all of the authors on several formal measures of excellence before he [it?] was outed).

research offices and faculties turn this goal into reality through centres and programmes of excellence, which are in turn linked through networks like the Canadian “Networks of Centres of Excellence” or German “Clusters of Excellence” (Networks of Centres of Excellence of Canada 2015; OECD 2014). Funding agencies use “excellence to recognize excellence” (Nowotny 2014).

The academic funding environment, likewise, is saturated with this discourse. A study of the National Endowment for the Humanities (NEH) is entitled *Excellence and Equity* (S. Miller 2015). The Wellcome Trust, a large medical funder, has grants for “sustaining excellence” (“Sustaining Excellence Awards” 2016). The National Institutes for Health (NIH), the largest funder of civilian science in the United States, claims to fund “the best science by the best scientists” (Nicholson and Ioannidis 2012) and regularly supports “centres of excellence.” The University Grants Commission of India recently awarded fifteen institutions the title of “University with Potential for Excellence” (University Grants Commission 2016). In the United Kingdom, the “Research Excellence Framework” uses “excellence” as a means of channelling differential funding to departments and institutions. In Australia, the national review framework is known as “Excellence in Research for Australia” (ERA) (“Excellence in Research for Australia” 2016). In Germany, the Deutsche Forschungsgemeinschaft (DFG) supports its “Clusters of Excellence” through a long standing “Excellence Initiative” (OECD 2014).

As this range of examples suggests, “excellence,” as used by Universities and their funders, is a flexible term. It can describe alike the activities of the world's top research universities and its smallest liberal arts colleges. It applies to their teaching, research, and management. It can encompass simultaneously the work of their Synthetic Biologists and Urban Sociologists, Anglo-Saxonists and Concert Pianists. It defines their Centres for Excellence in Teaching and their Centers of Excellence for Mechanical Systems Innovation (“USC Center for Excellence in Teaching” 2016; The University of Tokyo Global Center of Excellence 2016). It defines success in academic endeavour from Montreal to Mumbai.

But what does “excellence” mean? How is this apparently ubiquitous quality of academic life defined? Is there a single standard for identifying “excellence”? Or is “excellence” defined on a discipline-by-discipline, or case-by-case basis? Does the search for “excellence,” its use to reward and punish individual institutions and researchers, and its utility as a criterion for the organisation of research help or hinder the actual production of research and scholarship in the world academy? Tertiary education enrolls approximately 32% of world’s student age population, and OECD countries spent on average 1.6% of their GDP on University-level teaching and research in 2015; the U.S. alone spent 2.7% or \$484 billion (*The Economist* 2015). Is “excellence” really the most efficient metric for distributing the resources available to the

world's scientists, teachers, and scholars? Does "excellence" live up to the expectations we place upon it? Is "excellence" excellent?

This article examines the utility of "excellence" as a means for organising, funding, and rewarding science and scholarship. It argues that academic research and teaching is *not* well served by this rhetoric. Nor, we argue, is it well served by the use of "excellence" to determine the distribution of resources and incentives to the world's researchers, teachers, and research institutions. While the rhetoric of "excellence" may seem in the current climate to be a natural method for determining which researchers, institutions, and projects should receive scarce resources, we demonstrate that it is neither as efficient nor as accurate as it may at first seem. As we shall show, indeed, a focus on "excellence" impedes rather than promotes scientific and scholarly activity: it at the same time discourages *both* the intellectual risk-taking required to make the most significant advances in paradigm-shifting research *and* the careful "Normal Science" (Kuhn and Hacking 2012) that allows us to consolidate our knowledge in the wake of such advances. It encourages researchers to engage in counterproductive conscious and unconscious gamesmanship. And it impoverishes science and scholarship by encouraging concentration rather than distribution of effort. The net result is science and scholarship that is less reliable, less accurate, and less durable than research assessed according to other criteria.

The article itself falls into three parts. In the first section, we discuss "excellence" as a rhetoric. Drawing on work by Michèle Lamont and others, we argue that "excellence" is less a discoverable quality than a linguistic interchange mechanism by which researchers compare heterogeneous sets of disciplinary practices. In the second section, we dig more deeply into the question of "excellence" as an assessment tool: we show how it distorts research practice while failing to provide a reliable means of distinguishing among competing projects, institutions, or people. In the final section, we consider what it might take to change our thinking on scarcity and "excellence." We consider alternative narratives for approaching the assessment of research activity, practitioners, and institutions and discuss ways of changing the "scarcity-thinking" that has led us to our current use of this fungible and unreliable term.

## What we talk about when we talk about "excellence"

In her book, *How Professors Think: Inside the Curious World of Academic Judgment*, Michèle Lamont opens by noting that "'excellence' is the holy grail of academic life" (Lamont 2009, 1). Yet, as she quickly moves to highlight, this "excellence is produced and defined in a multitude of sites and by an array of actors. It may look different when observed through the lenses of peer review, books that are read by generations of students, current articles published by 'top' journals, elections at national academies, or appointments at elite institutions" (3). Or as Jack

Stilgoe argues, “‘Excellence’ is an old-fashioned word appealing to an old-fashioned ideal. ‘Excellence’ tells us nothing about how important the science is and everything about who decides” (Stilgoe 2014).

This tallies with the work of others who have considered reforms to the review process in recent years. Kathleen Fitzpatrick, for instance, has also situated the crux of evaluation in the evaluator, not the evaluated. For, as Fitzpatrick notes,

“in using a human filtering system, the most important thing to have information about is less the data that is being filtered, than the human filter itself: who is making the decisions, and why. Thus, in a peer-to-peer review system, the critical activity is not the review of the texts being published, but the review of the reviewers.” (Fitzpatrick 2011, 38)

The challenge here is that it is not possible to conduct a total “review of the reviewers” without some reference to the evaluated material. It may be possible to query the *conduct* of reviewers against another set of disciplinary norms (i.e. are the reviewers acting in good faith? Have they provided a useful report? Do they know the field as normatively defined?); but to review reviewers’ judgment of a specific work also requires an external evaluation of the work itself. In each case the challenge comes from a type of circularity in which a *pre-shared* evaluative culture must exist in order to pass judgment: the “shared standards” of which Lamont writes (2009, 4).

Yet despite the anti-foundationalism of such philosophies, there remains a pressing need, in Lamont’s view, to ensure that “peer review processes [... are] themselves subject to further evaluation” (247). Calls for training in peer review practices as well as calls for greater transparency occur across disciplinary boundaries, but generally without addressing the differences in *practice* that occur on either side of those boundaries. Lamont suggests that current remedies to this problem—which mostly consist of changing the degrees of anonymity or the point at which review is conducted (filter first vs. post-filter) within different peer-review practices—are insufficient and constitute “imperfect safeguards.” Instead, she suggests, it is more important that members of peer-review communities should be educated “about how peer evaluation works’, avoiding the pitfalls of homophily (in which review processes merely re-inscribe value to work that exhibits similitude to the disciplinary norm) by re-framing the debate as a “micro-political process of collective decision making” that is “genuinely social” (246-247). As with most problems in scholarly communications, the challenge with peer review is therefore not technical, but social.

As Lamont shows, then, "excellence" is a pluralised construct that is specific to each disciplinary environment. Yet even the most obvious solution to this challenge—interdisciplinary diversity of evaluators—only leads to further problems. For the differences in practice of review and perceptions of excellence across disciplinary boundaries, combined with a lack of appreciation that these differences exist, makes it difficult to reach consensus within diverse pools of reviewers. This is because, as Andy Stirling has noted, "it is difficult indeed to contemplate any single general index of diversity that could aggregate properties [...] in a uniquely robust fashion" (Stirling 2007). If diversity itself cannot easily be collapsed onto a single measurable vector then there is little hope of aggregating diverse senses of "excellence" into a coherent and universal framework.

The fact that "excellence" is a pluralised concept, used differently by different communities, places it under the category of what Susan Leigh Star terms a "boundary object" (Star 1989). In subsequent work, Star affirmed that boundary objects are those that possess three characteristics in the dynamics of their use:

1. The object [...] resides between social worlds (or communities of practice) where it is ill structured.
2. When necessary, the object is worked on by local groups who maintain its vaguer identity as a common object, while making it more specific, more tailored to local use within a social world, and therefore [only potentially] useful for work that is NOT interdisciplinary.
3. Groups that are cooperating without consensus tack back-and-forth between both forms of the object. (Star 2010)

Each of these elements applies to some extent to "excellence," though, as we shall see, not without caveats. Certainly, "excellence" resides between different communities and is ill-structured/defined in each context. Local groups and disciplines have their own more specific contexts of excellence: those in the biological sciences may treat some aspects of performance as "excellent" (e.g. number of publications, author position, citations counts), while failing to recognise aspects considered equally "excellent" by English professors (large word counts, single authorship, publication or review in popular literary magazines and journals)(O'Donnell 2015). Finally, as we shall see, it is clear that evaluative cultures are operating without even internal consensus beyond a few broad categories of performance.

Considering "excellence" in terms of a boundary object is worthwhile because it is tempting to argue that such concepts of value, even if they are ungrounded and unshared, can be used pragmatically to foster consensus. This is the point of Ludwig Wittgenstein's famous "beetle in a

box” metaphor, which he uses to exemplify the “private language argument” (Wittgenstein 2001, sec. 293). For Wittgenstein, the question of unique non-communicable epistemic knowledge (such as pain experience), should actually be framed in terms of public, pragmatic language games/contexts. If we each have an object in a box that is called a “beetle,” but none of us can see each other’s “beetles,” he argues, then the important thing is not what the objects in our boxes actually are but rather how we negotiate and use the term socially to engender intersubjective understanding or action. In such cases, “if we construe the grammar of the expression of sensation on the model of ‘object and designation’, the object drops out of consideration as irrelevant” and designation is all that matters.

We might therefore productively ask: even if “excellence” is a concept that carries little or no information content, either within communities or across them, might it nonetheless be useful as a “beetle”? That is as a carrier of interpretation or a set of social practices functioning as an expert system to convert intrinsic, qualitative, and non-communicable assessment into a form that allows performance to be compared across disciplinary or other boundaries? It might, indeed, be useful given the political necessity for research communities and institutions to present an (ostensibly) unified front to government and wider publics as a means of protecting their autonomy. Could “excellence” be, to speak bluntly, a linguistic signifier without any agreed upon referent whose value lies in an ability to capture cross-disciplinary value judgements and demonstrate the political desirability of public investment in research and research institutions?

In actual practice, as we shall see, it is not even useful in this way. Although, as its ubiquity suggests, “excellence” is used across disciplines to assert value judgements about otherwise incomparable scientific and scholarly endeavours, the concept itself mostly fails to capture the disciplinary qualities it claims to define. Because it lacks content, “excellence” serves in the broadest sense solely as an indicator of comparative success: that some thing, person, activity, or institution “better” or “more important” than some other (often otherwise incomparable) thing, person, activity, or institution—and, crucially, that it is, as a result, *more deserving of reward*. But this emphasis on reward, as Alfie Kohn and others have demonstrated, is itself often poisonous to the *actual* quality of the underlying activity (Kohn 1999). As we shall see in the following section, it encourages system-gaming and other undesirable behaviour (including at times, outright fraud) in those whose “excellence” is being evaluated and it de-incentivizes the kind of “normal” behaviour necessary for most successful scholarship and science (Kuhn and Hacking 2012; Kuhn 1963). While “excellence” appears to allow us to make comparative qualitative claims about work in different (sub-)disciplines, in other words, it does so without capturing essential features of the underlying work and by encouraging bad practice in individual researchers and teams that often works against the realisation of these essential features.

## Is “excellence” good for research?

Thus far, we have been arguing that “excellence” is primarily a rhetorical signalling device used to claim value across heterogeneous institutions, researchers, disciplines, and projects rather than a measure of intrinsic and objective worth. In some cases, as we shall see, the qualities of these projects can be compared in detail on other bases; in many—perhaps most—cases, they cannot. As we have argued, the claim that a research project, institution, or practitioner is “excellent” is little more than an assertion that that project, institution, or practitioner succeeds better on its own terms than some other project, institution, or practitioner succeeds on some other, usually largely incomparable, set of terms.

But what about these sets of “own terms”? How easy is it to define the “excellence” of a given project, institution, or practitioner on an intrinsic basis? Even if we leave aside the comparative aspect, are there formal criteria we can use to identify excellence in a *single* research instance on its own terms?

Research suggests that this is far harder than one might think. Academics, it turns out, appear to be particularly poor at recognising a given instance of “excellence” when they see it, or, if they think they do, getting others to agree with them. Their continued willingness to debate relative quality in these terms, moreover, creates a basis for extreme competition that has serious negative consequences.

## Do researchers recognise excellence when they see it?

The short answer is no. This can be seen most easily when different potential measures of excellence conflict in their assessment of a single paper, project, or individual. Adam Eyre-Walker and Nina Stoletzki, for example, conclude that scientists are poor at estimating the merit and impact of scientific work even after it has been published (2013). Post-publication assessment is prone to error and biased by the journal in which the paper is published. Predictions of future impact as measured by citation counts are also generally unreliable, both because scientists are not good at assessing merit consistently across multiple metrics and because the accumulation of citations is itself a highly stochastic process, such that two papers of similar merit measured on other bases can accumulate very different numbers of citations just by chance. Moreover, Wang et al. show that in terms of citation metrics the most novel work is systematically undervalued over the time frames that conventional measures use, including, for instance, the Journal Impact Factor that Eyre-Walker and Stoletzki suggest biases expert assessment (2016).

This is true even of work that can be shown to be successful by other measures. Campanario, Gans and Shepherd, and others, for example, have traced the rejection histories of Nobel and other prize winners, including for papers reporting on results for which they later won their recognition (Campanario 2009; Gans and Shepherd 1994; Azoulay, Zivin, and Manso Fall 2011, 527–528). Campanario and others have also reported on the initial rejection of papers that later went on to become among the more highly-cited in their fields and/or the journals that ultimately accepted them (Campanario 1993; Campanario 1996; Campanario 1995; Campanario and Acedo 2007; Siler, Lee, and Bero 2015; Nicholson and Ioannidis 2012; Calcagno et al. 2012). Yet others have found a generally poor relationship between high ratings in grant competitions and subsequent “productivity” as measured by high publication or citation counts (Costello 2010; Fang, Bowen, and Casadevall 2016; Lindner and Nakamura 2015; Meng 2016; Pagano 2006).

As this suggests, academics’ abilities to distinguish the “excellent” from the “not-excellent” do not correlate well with one another even within the same disciplinary environment (there tends to be greater agreement at the other end of the scale, distinguishing the “not acceptable” from the acceptable, though not necessarily “excellent”; see Cicchetti 1991; Weller 2001). In order to earn citations or win prizes for a rejected manuscript, after all, future Nobel prize winners need to begin by convincing a different journal (and its referees) to accept the work others previously found wanting. But this is not something that only future prize winners are good at: as Weller reported in the early years of this century, most (57%) rejected manuscripts were ultimately published; in the vast majority of cases (90%), these previously rejected articles were accepted on their second submission and, in the vast majority of these cases (also 90%), at a journal of similar prestige and circulation (Weller 2001). While these statistics have almost certainly changed in the last few years with changes in the demographics of submission and, especially, the development of venues that focus on the publication of “sound science” (Science 2016), the basic sense that journal peer review is a gatekeeper that is frequently circumvented remains.

Articles that are initially rejected and then go on to be published to great acclaim and/or in journals of a similar or higher ranking represent what are in essence false negatives in our ability to assess “excellence.” They are also evidence of terrible inefficiency. The rejection of papers that are subsequently published with little or no revision at journals of similar rank increases the costs for everyone involved without any countervailing improvement in quality. In addition to multiplying the systemic cost of refereeing and editorial management by the number of resubmissions, such articles also present an opportunity cost to their authors through lost chances to claim priority for discoveries, for example, or, even more commonly, lost opportunities for citation and influence (Gans and Shepherd 1994; Campanario 2009; Brembs 2015; Psych Filedrawer 2016; Şekercioğlu 2013).



More worryingly, there is also considerable evidence of false positives in the review process—that is to say that submissions that are judged to meet the standards of “excellence” required by one funding agency, journal, or institution, but do worse when measured against other or subsequent metrics. In somewhat controversial work, Peters and Ceci submitted papers in slightly disguised form to journals that had previously accepted them for publication (Peters and Ceci 1982; see Weller 2001 for a critique). Only 8% overall of these resubmissions were explicitly detected by the editors or reviewers to which they were assigned. Of the resubmissions that were not explicitly detected, approximately 90% were ultimately rejected for methodological and/or other reasons by the same journals that had previously published them; they were rejected, in other words, for being insufficiently “excellent” now by journals that had previously decided they were “excellent” enough to enter the literature.

When it comes to funding, a similar pattern of false positives may pertain: a study by Nicholson and Ioannidis suggests that highly cited authors are *less* likely to head major biomedical research grants than less-frequently-cited but socially-better-connected authors who are associated with granting agency study groups and review panels (Nicholson and Ioannidis 2012). Fang, Bowen and Casadevall have discovered that “the percentile scores awarded by peer review panels” at the NIH correlated “poorly” with “productivity as measured by citations of grant-supported publications” (Fang, Bowen, and Casadevall 2016). These suggest a bias towards conformance and social connectedness over innovation in funding decisions in a world in which success rates are as low as 10%. It also provides further evidence of funding-agency bias against disruptively innovative work noted by many researchers over the years (Siler, Lee, and Bero 2015; Kuhn and Hacking 2012; Campanario 1995; Campanario 1996; Campanario 1993; Campanario 2009; Costello 2010; Ioannidis et al. 2014).

To the extent that the above are evidence of inefficiencies in the system, some might argue that individual problems in determining “excellence” in these cases are resolved in the longer term and over large sample groups. Of course, these examples only show work for which multiple measures of “excellence” can be compared: given their unreliability, this suggests that work that is not measured more than once may be unjustly suppressed or unjustly published. Either way, however, they represent honest difference of opinion. The same cannot be said, however, of actual fraud and outright errors.

### **Fraud, error and lies**

As various studies have concluded, reported instances of both fraud and error (as measured through retractions) are on the rise (Andrade 2016; Steen 2011; Fang, Steen, and Casadevall 2012; Chen et al. 2013; Grieneisen and Zhang 2012; Claxton 2005; Dobbs 2006; Yong 2012b).

This is particularly true at higher prestige journals (Belluz 2016; Siler, Lee, and Bero 2015; Resnik, Wager, and Kissling 2015-7). If we add to this list of (potentially) “false positives” studies that cannot be replicated, we increase the number of papers that meet one measure of “excellence” (i.e. passing peer review, often at “top” journals) while failing others (i.e. being accurate and reproducible, and/or non-fraudulent) (Hill and Pitt 2014; Open Science Collaboration 2015; Rehman 2013; Yong 2012b; Burman, Reed, and Alm 2010; Chang and Li 2015; Dean 1989; Bem 2011; Goldacre 2011; Resnik and Dinse 2013; Lehrer 2010). As we shall see below, however, it is the very focus on “excellence” in publication that encourages the submission of fraudulent, erroneous, and irreproducible papers, and that also works to prevent the publication of reproduction studies that can identify such problems.

Erroneous, and especially fraudulent or irreproducible papers are interesting because they represent a failure of *both* our ability to identify and predict actual “excellence” *and* the incentive system we use to encourage scientists and scholars to aim for this same quality, illustrating just how pernicious our collective focus on “excellence” can be. As Fang, Steen, and Casadevall have shown, the majority of retracted papers are withdrawn for reasons of misconduct including fraud, duplicate publication, or plagiarism (67.4%), rather than error (21.3%) (2012; cf. Steen 2011 for which the later article represents a correction). But even these figures may under-represent the true incidence of misconduct. As focus groups and surveys conducted by various researchers have demonstrated, error itself can represent misconduct in the form of a (semi-)deliberate strategy for ensuring quick and/or numerous publications by “cutting a little corner” in order to get a paper out before others or to get a larger grant,... [or] because... [a researcher] needed more publications that year” (Anderson et al. 2007, 457–458; see also Fanelli 2009; Chubb and Watermeyer 2016; Tijdink, Verbeke, and Smulders 2014). In one small sample of detailed surveys, for example, Fanelli showed that while only a small percentage of scientists (1.97% n=7) admitted to fabricating, falsifying, or modifying data, a much larger percentage claimed to have seen others engaging in similarly outright fraudulent activity (14.12% n=12). Furthermore, even larger percentages had engaged in (33.7%) or seen others engage in (72%) questionable research described using less negatively loaded language (Fanelli 2009; the percentage of scientists admitting to explicit misconduct is considerably higher [15%] in Tijdink, Verbeke, and Smulders 2014). As Fanelli concludes: “Considering that these surveys ask sensitive questions and have other limitations, it appears likely that this is a conservative estimate of the true prevalence of scientific misconduct” (2009, 9)—a conclusion very strongly supported by the anecdotal admissions of Anderson et al.’s focus groups.

The drive for excellent performance in the eyes of assessors is shown even more starkly in work by Chubb and Watermeyer (2016). In structured interviews, academics in Australia and the United Kingdom admitted to outright lies in the claims of broader impacts made in research

proposals. As the authors note: “[h]aving to sensationalize and embellish impact claims was seen to have become a normalized and necessary, if regretful, aspect of academic culture and arguably par for the course in applying for competitive research funds” (6). Quoting an interviewee, they continue, “[i]f you can find me a single academic who hasn’t had to bullshit or bluff or lie or embellish in order to get grants, then I will find you an academic who is in trouble with his [sic] Head of Department” (6; “[sic]” as in Chubb and Watermeyer). Here we see how a competitive requirement, perceived or real, for “excellence,” in combination with a lack of belief in the ability of assessors to detect falsehoods, leads to a conception of “excellence” as pure performance: a concept defined by what you can get away with claiming in order to suggest (rather than actually accomplish) “excellence.”

What is striking about these behaviours, of course, is that they are unrelated to (and to a great extent perhaps even incompatible with or opposed to) the actual qualities funders, governments, and the research community are ostensibly using “excellence” to identify. No agency, ministry, or university research office intentionally uses “excellence” as shorthand for “able to embellish results or importance convincingly,” even as the researchers adjudicated under this system report this as a primary criterion for success. Whether it occurs through fraud, cutting corners, or embellishing sections of grant proposals, the performance of “excellence” is commonly justified as being necessary for survival, suggesting a cognitive and cultural dissonance between those aspects of their work that the performers feel is essential and those aspects they feel they must emphasise, overstate, embellish, or fabricate in order to appear more “excellent” than their competitors. The increasing evidence that fraud and corner-cutting are a problem at the core of the research process suggests that the pressure for these performances of “excellence” is not restricted to stages that do not matter. As Kohn argues, reward-motivation affects scientific creativity (the ability to “break out of the fixed pattern of behavior that had succeeded in producing rewards... before”) as much as it does evidence gathering or the inflation of results (1999, 44; see also Azoulay, Zivin, and Manso Fall 2011; Tian and Wang 2011; Lerner and Wulf 2006).

### **The “winner’s curse”: Competition for scarce resources and the performance of “excellence”**

So why do researchers engage in this kind of dubious activity? Clearly for both Chubb and Watermeyer’s interviewees, as well as those identified as having committed scientific fraud (from whom we have testimony, the motivation, if not always the justification), it is competition for scarce resources, whether funding, positions, or community prestige. Of course this is not a new issue (Smith 2006). Taking time away from his work on the difference machine, Charles Babbage published an analysis of what he saw as the four main kinds of scientific frauds in an

1830 polemic, *Reflections on the Decline of Science in England: And on Some of Its Causes*. These included the self-explanatory “hoaxing” and “forging,” in addition to “trimming” (“clipping off little bits here and there from those observations which differ most in excess from the mean and in sticking them on to those which are too small”) and “cooking” (“an art of various forms, the object of which is to give ordinary observations the appearance and character of those of the highest degree of accuracy”) (Babbage 1830, 178; see Zankl 2003; and Secord 2015 for a discussion).

The motivation for these frauds, then as now, involves prestige and competition for resources. Babbage’s typology of fraudulent science was but a minor chapter in a book otherwise mostly concerned with the internal politics of the Royal Society. Babbage attributed the decline he saw in English science to the lack of attention and professional opportunities available to potential scientists, and, as a result, was keenly sensitive to questions of credit and its importance in determining rank and authority. Indeed, as Casadevall and Fang remind us, “[s]ince Newton, science has changed a great deal, but this basic fact has not. Credit for work done is still the currency of science.... Since the earliest days of science, bragging rights to a discovery have gone to the person who first reports it” (Casadevall and Fang 2012, 13). The prestige of first discovery was always a scarce resource. Now that that prestige is measured through the scarce resource of authorship in the “right journals” and is coupled ever more strongly to the further scarce resources of career advancement and grant funding it should not be a surprise that the competition for those markers has become steadily stronger and the performance of “excellence” has become more marked.

In the course of the last three quarters of a century, the value of this basic currency of credit has only increased in response to the increasing cost and competitiveness of the scientific endeavour (on the roughly 3000% increase in constant dollars in government funding for science in the U.S. since the early 1950s, see Stephan 2014, 114–116; de Solla Price 1975 estimates that scientific publication has grown historically at a rate of about 4.7% per year; this rate was confirmed for the period 1990–2007 by Larsen and von Ins 2010; interestingly, Ioannidis, Boyack, and Klavans 2014 estimate that <1% of active scientists publish every year). As Casadevall and Fang argue:

The winner-take-all aspect of the priority rule has its drawbacks, however. It can encourage secrecy, sloppy practices, dishonesty and an excessive emphasis on surrogate measures of scientific quality, such as publication in high-impact journals. The editors of the journal *Nature* have recently exhorted scientists to take greater care in their work, citing poor reproducibility of published findings, errors in figures, improper controls,

incomplete descriptions of methods and unsuitable statistical analyses as evidence of increasing sloppiness....

As competition over reduced funding has increased markedly, these disadvantages of the priority rule may have begun to outweigh its benefits. Success rates for scientists applying for National Institutes of Health funding have recently reached an all-time low. As a result, we have seen a steep rise in unhealthy competition among scientists, accompanied by a dramatic proliferation in the number of scientific publications retracted because of fraud or error. Recent scandals in science are reminiscent of the doping problems in sports, in which disproportionately rich rewards going to winners has fostered cheating. (2012, 13)

If scandals such as fraudulent articles were the only way in which this overwhelming competitive focus on “excellence” hurt research, it would be bad enough. But the intense competition also leads to “the Matthew effect”—i.e. the disproportionate accrual of incentives to those researchers and institutions that best demonstrate “excellence”—creating distortions throughout the research cycle, even for work that is not fraudulent or the result of misconduct (Bishop 2013; as its etymology implies, the “Matthew effect” predates today’s hypercompetition, see Merton 1968; Merton 1988)<sup>2</sup>: as we have seen above, this competition and these rewards reduce creativity; encourage gamesmanship (and concomitant defensive conservatism on the part of review panels) in granting competitions; create a bias towards ostensibly novel (though largely non-disruptive), positive, and even inflated results on the part of authors and editors; and they discourage the pursuit and publication of replication studies, even when these call into serious question important results in the field.

N. S. Young, Ioannidis, and Al-Ubaydli have described the effect of this incentive system as an example of the “winner’s curse,” a concept from auction theory which explains why, “under certain conditions, the bidder who wins tends to have overpaid” (2008, 1418). While, in economic theory, the average of all bidders’ estimates in an auction would normally be assumed to approximate the value of the object being sold, winners win auctions by outbidding their opponents—that is to say by paying more than an average bidder would say the object is worth. As they go on to argue:

An analogy can be applied to scientific publications. As with individual bidders in an auction, the average result from multiple studies yields a reasonable estimate of a “true”

---

<sup>2</sup> The Matthew Effect is derived from *Matthew* 13:12: ‘For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath.’

relationship. However the more extreme, spectacular results (the largest treatment effects, the strongest associations, or the most unusually novel and exciting biological stories) may be preferentially published. Journals serve as intermediaries and may suffer minimal immediate consequences for errors of over- or mis-estimation, but it is the consumers of these laboratory and clinical results (other expert scientists; trainees choosing fields of endeavour; physicians and their patients; funding agencies; the media) who are “cursed” if these results are severely exaggerated—overvalued and unrepresentative of the true outcomes of many similar experiments. (N. S. Young, Ioannidis, and Al-Ubaydli 2008, 1418)

In other words, understood from the perspective of auction theory, scientists “buy” career good in the form of pages, positions, funding (all of which result in CV lines) from the most prestigious (and hence scarce) journals, presses, and grant competitions by consciously or unconsciously “overbidding” their results—from emphasising their novelty or human interest (Gonon et al. 2012; Vinkers, Tijdkink, and Otte 2015; Atkin 2002) to, in Babbage’s terms “trimming,” “cooking,” and, in a few cases, “hoaxing” and “forging” (Fanelli 2009; Tijdkink, Verbeke, and Smulders 2014; Anderson et al. 2007; Chubb and Watermeyer 2016). In what is under such conditions in essence a sellers’ market, the cost of demonstrating ever increasing “excellence” can be increasingly impoverished science and scholarship. The greater the degree of competition, the greater the potential winners’ curse.

### Positive bias and the decline effect

The idea of a “sellers’ market” in which researchers “perform excellence” is supported by the well known bias towards positive results in scientific publication (e.g. Psych Filedrawer 2016; Rothstein 2014; Kennedy 2004; S. S. Young and Bang 2004; Dickersin et al. 1987; Dickersin 2005; Sterling 1959; Bertamini and Munafò 2012). Thus, for example, Fanelli demonstrated a 22% growth between 1990 and 2007 in the “frequency of papers that, having declared to have ‘tested’ a hypothesis, reported a positive support for it” (Fanelli 2011). This is all the more remarkable, given that the late 1980s were themselves not a halcyon period of unbiased science: in an 1987 study of 271 unpublished and 1041 published trials, Dickersin et al found that 14% of unpublished and 55% of published trials favoured the experimental therapy (1987). As Fanelli argues: “Methodological artefacts cannot explain away these patterns, which support the hypotheses that research is becoming less pioneering and/or that the objectivity with which results are produced and published is decreasing.” Or as N. S. Young et al. suggest, “the general paucity in the literature of negative data” is such that “[i]n some fields, almost all published studies show formally significant results so that statistical significance no longer appears discriminating” (2008, 1419).

Another artifact of this “winner’s curse” is the “decline effect,” or the tendency for the strength of evidence for a particular finding to decline over time from that stated on its first publication (Brembs, Button, and Munafò 2013; Open Science Collaboration 2015; Groppe 2015; Schooler 2011; Gonon et al. 2012). While this effect is also well-known, Brembs et al. have recently shown that its presence is significantly positively correlated with journal prestige as measured by Impact Factor: early papers appearing in high prestige journals report larger effects than subsequent studies using smaller samples (2013, fig. 1b and 1c).

### The bias against replication

Finally, there is the bias against the publication of replication studies in disciplines where such patterns make scientific sense. As Lehrer (who ironically withdrew two books from publication because of fabricated quotations—a case of “over-paying” in the intense competition for placement on the *New York Times* bestseller list) argues

The test of replicability, as it’s known, is the foundation of modern research. Replicability is how the community enforces itself. It’s a safeguard for the creep of subjectivity. Most of the time, scientists know what results they want, and that can influence the results they get. The premise of replicability is that the scientific community can correct for these flaws.... For many scientists, the effect is especially troubling because of what it exposes about the scientific process. If replication is what separates the rigor of science from the squishiness of pseudoscience, where do we put... rigorously validated findings that can no longer be proved? (Lehrer 2010)

Despite this, however, science is facing a replication crisis due to the low incentives for replication work. As Nosek et al. note

Publishing norms emphasize novel, positive results. As such, disciplinary incentives encourage design, analysis, and reporting decisions that elicit positive results and ignore negative results. Prior reports demonstrate how these incentives inflate the rate of false effects in published science. When incentives favor novelty over replication, false results persist in the literature unchallenged, reducing efficiency in knowledge accumulation. Previous suggestions to address this problem are unlikely to be effective. For example, a journal of negative results publishes otherwise unpublishable reports. This enshrines the low status of the journal and its content. (2012)

Even more remarkable, however, is the bias that exists against publishing such studies *even when they invalidate the original, often high profile, result* and, as a result, might be expected to have an increased profile themselves (Aldhous 2016; Wilson 2011; Nosek, Spies, and Motyl

2012; Yong 2012b; Yong 2012a; Goldacre 2011; for a view from the other side of replication, see Bissell 2013). This is in part, a function of publishing economics: commercial journals (including “non-commercial” journals whose income is used to support scholarly and scientific societies) earn money from subscription, access, and reprint fees (Lundh et al. 2010); high profile results and a high prestige reflected by a high Impact Factor help maintain the demand for these journals and hence ensure both a continuing stream of interesting new material and a steady or rising income for the journal as a whole (Lundh et al. 2010; Marcovitch 2010; Lawrence 2007; Munafò, Stothart, and Flint 2009). But it is also a result of the incentive system that guides authors as well, as Wilson notes:

[M]ajor journals simply won't publish replications. This is a real problem: in this age of Research Excellence Frameworks and other assessments, the pressure is on people to publish in high impact journals. Careful replication of controversial results is therefore good science but bad research strategy under these pressures, so these replications are unlikely to ever get run. Even when they do get run, they don't get published, further reducing the incentive to run these studies next time. The field is left with a series of "exciting" results dangling in mid-air, connected only to other studies run in the same lab. (2011)

Or as Rothstein argues “The consequences of this problem include the danger that readers and reviewers will reach the wrong conclusion about what the evidence shows, leading at times to the use of unsafe or ineffective treatments” (2014).

## Homophily

Thus far, we have been discussing the negative impact of “excellence” largely in terms of its effect on the practice and result of professional researchers. There is, however, another effect of the drive for excellence: a restriction in the range of research and scholarship performed and the impact such research and scholarship has on the larger population. Although “excellence” is commonly presented as the most fair or efficient way of distributing scarce resources (Sewitz 2014), it in fact can have an impoverishing effect on the very practices that it seeks to encourage. A funding programme that looks to improve a nation’s research capacity by differentially rewarding “excellence” can have the paradoxical effect of reducing this capacity by underfunding the very forms of “normal” work that make science function (Kuhn and Hacking 2012). A programme that seeks to reward Humanists, likewise, by focussing on output paradoxically reduces the impact of these same disciplines by encouraging researchers to focus on their professional peers rather than broader cultural audiences (Readings 1996). A programme of concentration on the “best” academics, in other words, can have the effect of focussing attention



on problems and approaches that can be most easily demonstrated to be “excellent,” rather than those that are the most important or potentially impactful.

Homophily is in some senses a variant on Merton’s “Matthew effect,” discussed above (Merton 1968); it is also extremely harmful to the promotion of diversity in science and scholarship. Given the strong evidence that there is systemic bias within the institutions of research against women, under-represented ethnic groups, non-traditional centres of scholarship, and other disadvantaged groups (for a forthright admission of this bias with regard to non-traditional centres of scholarship, see Goodrich 1945), it follows that resources will not be allocated optimally, because the qualities of work from underrepresented groups will not always be recognised (King et al. 2015; King et al. 2014; University of Arizona Commission on the Status of Women 2015; O’Connor and O’Hagan 2015). There is a clear case to answer that, absent substantial corrective measures, a focus on excellence will not achieve optimal outcomes on this area.

This is a variant on an old argument, that existing power structures – those populated by those who exemplify “excellence” – tend to conservatism in their processes of evaluation. It underpins the calls to reassess the focus of mainstream scholarship, whether this is “great men” history, through “the canon” in literary studies, to the focus of disease research on the ills of rich American men. As Barbara Herrnstein Smith says with respect to literary evaluation:

...[a work that “endures”] will also begin to perform certain characteristic cultural functions by virtue of the very fact that it *has* endured...In these ways, the canonical work begins increasingly not merely to *survive within* but to *shape and create* the culture in which its value is produced and transmitted and, for that very reason, to perpetuate the conditions of its own flourishing. (Herrnstein Smith 1988 emphasis in the original)

That work, and those people, who are considered “excellent” will always be evaluated, like the canon that shapes the culture that transmits it, on a conservative basis. Whether viewed as a question of power and justice or simply as an issue of lost opportunities for diversity in the cultural co-production of knowledge, “excellence” will always necessarily be backwards looking, an evaluative process by institutions and individuals co-produced by a contingent historical context.

Merton’s concerns went deeper than questions of power and productivity, however. He viewed diversity of disciplinary perspectives as an important component of rigor (Merton 1973). It might well be argued that if diversity is important for good research, and that rhetorics of “excellence” damage diversity, then competition focussed around an “excellence” discourse risks

damaging the rigor of research itself. If diversity of perspectives is a strength of the critical pursuit of science and scholarship, then the pursuit of “excellence,” with its effect of increasing the concentration of resources in the hands of a few strong “performers,” is not simply unhelpful in determining the allocation of scarce resources: it is damaging to the very core of science and scholarship.

### Alternative Narratives: Working for change

This brings us, then, to the question of what is to be done. In the first section we showed that “excellence” is clearly not a concept with shared meaning across even sub-sections of the research community. In the second section we demonstrated how the pursuit of such “excellence,” applied within the context of specific disciplinary practice, leads to performative behaviour which reduces diversity, encourages questionable practices—including, at its worst, outright fraud and deception—while failing to support a range of desired outcomes, including rigor, creativity, and diversity of approach and opinion. We demonstrated how “excellence” as a social construct within a discipline is inherently conservative, looking back to the priorities of those people who and those outputs that have had the label conferred upon them in the past and, in many cases, are most embedded in the social fabric and expectations of a given discipline. And we showed how this in turn reinforces the re-performative aspect and draws resources and effort away from underpinning work such as replication.

This focus on how the “traditionally excellent” drives sub-optimal practices is not new. Indeed, many of the studies we review above have been driven explicitly or implicitly by an interest in the impact of “excellence” on encouraging poor research practice. There are, however, two new elements in our analysis.

The first is that we diagnose the problem not as the pursuit of quality *per se*, but as the concentration of attention (and resources) on the intense competition to make it into the top few percent—it is not “excellence” or its pursuit that is the problem; it is the *concentration on only the excellent*.

The second is in seeing this narrative as *internal*. The implicit claim in much of the work we have cited (and perhaps especially in the interviews and surveys in which researchers justify their “performance”) is that this competition and concentration is imposed by (semi-)outsiders: funders, governments, university administrators, a suspicious general public—a “them” whom researchers are forced to placate if they wish to gain the resources they need to pursue their “real” work. We, in contrast, trace the origins of this problem back inside the community of

researchers itself. It is the result of assumptions researchers bring with them about the nature of competition and their claimed ability to distinguish “excellence” from everything else.

Addressing this problem will require us to recognise, therefore, that it is inherent to the current system—a system in which resources are highly concentrated and attention is focussed at “the top” and on “the best.” But that raises the question of whether there is an alternative? Surely with limited resources, competition is the best—or perhaps only—approach to the distribution of scarce resources? This next section will explore the question of the evidence for scarcity and the necessity for competition and ranking as a distribution tool.

### Questioning the narrative

The narrative of “scarce resources” is a powerful one and one that dominates debate amongst researchers. But in some senses it is peculiar that it should have such force at this particular moment. Globally, investment in research and scholarship is at historically high levels (on the roughly 3000% increase in constant dollars in government funding for science in the U.S. since the early 1950s, see Stephan 2014, 114–116; de Solla Price 1975 estimates that scientific publication has grown historically at a rate of about 4.7% per year; this rate was confirmed for the period 1990-2007 by Larsen and von Ins 2010; interestingly, Ioannidis, Boyack, and Klavans 2014 estimate that <1% of active scientists publish every year; see also World Bank 2016 for research expenditures as a % of GDP). Indeed, senior members of the U.S. biomedical research establishment stated in 2014 that it was precisely the “longstanding assumption that the biomedical research system in the United States will expand indefinitely at a substantial rate” that was “the root cause of the widespread malaise” characterised by hyper-competition and an oversupply of junior researchers with limited opportunities for permanent research positions (Alberts et al. 2014).

Paradoxically, then, it appears that it is this very increase in funding (with the concomitant rise in the number of researchers) that has led to our sense of scarcity: the increase in forage has brought about an increase in fauna. But it has also brought about our sense that these “limited” resources must be *concentrated* on the “excellent” rather than distributed more broadly. This leads to a cycle in which those who control the resources define and exemplify the characteristics whose performance leads to the awarding of those same resources (Stilgoe 2014) leading to further concentration, further performance, and, according to our argument, damage to the research communities and processes they are supposed to support: hypercompetition for resources damages the herd. This cycle also solidifies existing issues of representation and equity arguably driving the research establishment further away from contact with wider publics.

To tackle this we need plausible counter-narratives to arise from within our research communities. Broadly speaking, these narratives will take two forms. The first is a move away from “excellence” as a boundary object and narrative towards more pluralistic and objective (or at least shared) conceptions of the qualities of research. The second will address the issue of concentration-driven hypercompetition by arguing for the redistribution of research resources. This second narrative is more challenging, as it is also a frontal attack on decades of policy and the power structures that have driven that agenda. But we will deal with each of these narratives in turn below.

### **Establishing a pluralistic counter narrative**

Our first narrative requires us to recognise that “excellence” is not the only (or even necessarily easiest) method of distributing resources and conducting research.

In fact, more pluralistic views on both the qualities of research that matter and the qualities of those qualified to contribute to research are already seen in two scholarly contexts: very large projects and new disciplinary formations. Large projects such as the Human Genome project and the Large Hadron Collider require coordination, technical specialisations, and funding much greater than that required by smaller research groupings. In order to get this funding, Nielsen shows, such megaprojects are forced to be far more open and transparent in their use and distribution of resources and opportunity than more traditional, smaller, projects: if they were not, referees would refuse to fund them due to their impact on the overall funding available for research in the discipline (2012, 109). Similarly new disciplinary formations, such as those created by the recent rise of the Digital Humanities and Bioinformatics, also demonstrate the extent to which a broader distribution of resources is possible and salutary. Such disciplines are generally founded in a collision of disciplinary technologies and cultures, particularly, in recent decades, the interaction of computation with traditional disciplines (though as the rise of “theory” in literary studies in the 1970s through 1990s, demonstrates, the colliding fields do not always require a technological component). The result of their formation, despite, usually, initial resistance from practitioners of the “traditional” disciplines undergoing expansion, is a broader sense of disciplinary opportunity and, in many cases, more resources.

Clay Shirky, looking particularly at the digital revolution, provides a way of understanding why this resistance occurs and why it is ultimately unsuccessful in holding back the development of new disciplinary approaches. In particular, he shows, the sense that information overwhelms previous disciplinary practice has to do with the extent to which new information technologies and new dissemination networks always create an increased abundance of content, overwhelming previous ways of understanding and discriminating among new material: or as he puts it, “it’s not information overload, it’s filter failure” (Shirky 2010). In fact, we would argue,

“discovery deficit” (Neylon 2010) might be a better term, given the extent to which “filter failure” is understood by researchers to imply a need to *strengthen* traditional (and as we have argued, conservative) methods of identifying “excellence,” a concentrative approach almost totally at odds with Shirky’s point. Regardless of the term used, however, the important thing for our argument is that this resistance invariably plays itself out during the development of new interdisciplinary advances in a field as an argument of “excellence”: critics of Bioinformatics, of the Digital Humanities, or even a pre-digital example like Critical Theory, do not generally argue that the new formations provide too rich an experience (though complaints of *ars longa vita breva* extend back to Hippocrates, their focus tends to be on the size of the traditional body of learning, not the expansion of a field); rather, critics of disciplinary expansion tend to argue that the new practices are less “excellent” than the old: less learned, less careful, less critical, or less inciteful, less methodologically sound—precisely the qualities of “excellence,” in other words, that would deny practitioners of the newly expanded domain access to funding, journal space, and positions to the detriment of the old (see also Stilgoe 2014 who emphasises the extent to which such rhetorics enforce the preservation of the status quo).

In practice, however, such criticism is always more Canutian than effective. “Resistance to theory,” or hostility to the meta-study of method and interpretation in literary studies (De Man 1982), was largely over by the beginning of the twenty-first century (see Wikipedia contributors 2016; of course one can’t have a “theory war” without an occasional post-armistice tussle: Jay 2014; Flaherty 2014; Bauerlein 2014; J. H. Miller 1987 was in any event probably premature in declaring Theory’s “triumph” by 1986). A recent broadside against the Digital Humanities in the *Los Angeles Review of Books* has the air of a (repetitive) rearguard action against a programme of research that has already become deeply entrenched in current disciplinary practice (Allington, Brouillete, and Columbia 2016; cf. Marche 2012; O’Donnell 2012; Pannapacker 2011b; Pannapacker 2012; Pannapacker 2011a). Both of these lines of attack have parallels with complaints from the 1980s about the Human Genome project (see Roberts 2001), while, more recently, criticism of the ENCODE project, or Bioinformatics more generally, could have been written as easily, *mutatis mutandis*, about practitioners of the Digital Humanities (on the ENCODE “functional gene” controversy, see Graur et al. 2013; comment and response in Nature 2013; Birney 2012; Birney 2016; Boyle 2013; on attitudes towards Bioinformatics more generally, see, especially the comments: Edwards 2015; Lowe 2013; Multiple contributors 2013; Brown 22 May, 2016; Rogan and Zou 2013).

These complaints ultimately fail to stop the rise of the new field, not because they are always without merit in specific cases (as Edwards 2015 notes, all specific examples of research are open to criticism, including those in “new” domains), but because the expansion of a given discipline or field of research through the addition of new methods and people is ultimately

never intellectually impoverishing. The new demands and increased number of researchers *can* thin out the resources available for all, contributing to a sense of scarcity in the original fields (though they can also bring an excitement that increases the pie, see, for example, in the case of the Digital Humanities, Wankel 2013, 8). But as Shirky, Nielsen, and others have shown, they also contribute new approaches and new problems to the research enterprise, solving old problems and introducing new ones in much the same way, Kuhn argues, that disruptive research creates new opportunities for the kind of “normal science” most researchers are engaged in most of the time (Kuhn and Hacking 2012; Shirky 2010; Nielsen 2012).

Nielsen’s *Reinventing Discovery* provides a compelling presentation of how an even wider distribution of attention and resources can improve research outcomes over the efforts of a more narrow, traditionally determined “excellence” through the application of “micro-expertise,” a form of cognitive surplus that occurs when open collaboration is emphasised over the concentration of data and resources in the hands of a few (2012). Thus, for example, Nielsen shows how the Polymath project, led by Tim Gowers (recipient of the Fields Medal, a traditional marker of mathematical “excellence”), used the collaborative efforts of professional mathematicians (including a second Fields medal winner) and interested amateurs to develop and solve complex problems (1-2). The same approach has been used in playing chess games (15-18), identifying galaxies (5-6 and *passim*), and, perhaps most-interestingly, as an amateur-instigated meta-study of the causes of migraines (91-93 and *passim*). It is also, of course, the engine that drives such now-familiar-but-once-novel crowd-sourced projects as the Wikipedia, Stack Exchange, or the development of Linux.

But Nielsen argues that this open, distributive approach is also increasingly the norm for successful projects at the other end of the institutional scale: massive, professionally organised and run, collaborative projects such as the Human Genome project; the Sloan Digital Sky Survey (SDSS); the Ocean Observatories Initiative; and the Allen Brain Initiative (Nielsen 2012, 108 and *passim*). As Nielsen notes, indeed, these large projects are, on the whole, far more likely to be open and distributive than smaller, team- or individual led teams:

It seems, then, that big scientific projects are more likely to make their data open than small projects. Why is that the case? Part of the explanation is political. Think about the SDSS. A typical small astronomy project may costs “only” a few tens or hundreds of thousands of dollars. That’s a lot of money, but it’s small change out of the billions of dollars our society spends on astronomy. If the people doing the experiment keep the data to themselves, it’s not a big loss to other astronomers. Further, those other astronomers aren’t in any position to complain for they too are keeping the data from their experiments secret. It’s a stable, uncooperative state of affairs. But the SDSS’s size

makes it special and different. It's so large that it consumes much of the entire world budget for astronomy. If the data is kept secret, then to astronomers outside the SDSS collaboration it's as though that entire chunk of money has simply disappeared from the astronomy budget. They have every reason to insist that the data be made open. And so, if large projects don't commit to at least partial openness, their applications for funding risk being shot down by people in the same field but outside the collaboration. This motivates big scientific projects to make their data at least partially open. (2012, 109)

Nielsen is writing particularly of access to data in this passage. But his observation about the undesirable effect of the concentration of resources on the practice of science in the case of such mega projects is also relevant to our argument here. The focus on "excellence" as a criterion for the distribution of funding and other career goods has the same effect of concentrating resources in an increasingly narrow group of winners. Instead of a *single* project consuming "much of the the entire world budget" for a given discipline, we instead see a small number of projects and researchers collecting resources to themselves. The net effect, however, is much the same: a few key players acquire the bulk of the resources and "it's as though that entire chunk of money has simply disappeared" from the resources remaining to all other practitioners. In an age in which networked science and large collaborative projects are demonstrating the power of distributed effort, concentrations of resources such as this in the hands of a few "excellent" researchers and institutions represents a huge opportunity cost. As Nielsen argues:

We're at a unique moment in history: for the first time we have an open-ended ability to build powerful new tools for thought. *We have an opportunity to change the way knowledge is constructed.* But the scientific community, which ought to be in the vanguard, is instead bringing up the rear, with most scientists clinging to their existing way of working, and failing to support those who seek a better way. As with the first open science revolution [i.e. the invention of journals in the seventeenth century], as a society we need to actively avert this tragedy of lost opportunity, by incentivizing and, where appropriate, compelling scientists to contribute in new ways. I believe that with hard work and dedication, we have a good chance of completely revolutionizing science.... When the history of the late twentieth century and early twenty[-]first centuries is written, we'll see this as *the* time in history when the world's information was transformed from an inert, passive state, and put into a unified system that brings that information alive.... And that change gives us the opportunity to restructure the way scientists think and work, and so to extend humanity's problem-solving ability. We are reinventing discovery, and the result will be a new era of networked science that speeds up discovery, not in one small corner of science, but across all of science. That

reinvention will deepen our understanding of how the universe works and help us address our most critical human problems. (2012, 206–207, emphasis as in original)

Nielsen's and Shirky's arguments operate at both the large scale of systems but also the individual scale of issues of diversity. Implicit in the notion of micro-expertise is the likelihood of valuable contributions coming from non-traditional, and non-certified sources. "Excellence" in its boundary-object form refers to the whole of a person or research output. The argument for greater *plurality* in assessment of *qualities* is in large part an argument for greater *granularity*—i.e. that this particular comment is valuable, despite this person not being an "expert."

Taking advantage of the opportunities that Nielsen identifies requires a substantial shift in attention away from those that exhibit (or perform) "excellence," a shift that he describes as the "redistribution of expert attention." But even if it is true, as Nielsen's examples would suggest, that a broader distribution of resources might unleash an underutilized capacity, the political argument for "taking" resources away from those with the patents and licences and "giving" it to those who can show less evidence of traditional "excellence" might be hard to make.

### Arguments for redistribution

In fact there have been arguments for more distributive research resourcing policy, but the voices are substantially fainter than the over-arching agenda towards concentration (see for example Nurse 2015). There are persistent arguments for and schemes designed to include some element of random (Bollen et al. 2014; Fang, Bowen, and Casadevall 2016; Health Research Council of New Zealand 2016) or equal distribution (Fortin and Currie 2013) of resources.

In practice, of course, the tendency in policy, at least in the traditional North Atlantic centres of research, has clearly been in the a non-distributive direction: for the concentration of resources on "top" institutions. The UK's Research Excellence Framework apportions untied funding of roughly £1.5B per year only to those institutions demonstrating work that is "internationally leading" or "internationally competitive" (Higher Education Funding Council for England 2016). Massive new research centres such as the Crick in London are intended to create a "critical mass" of "excellent" or "world-leading" research. In Canada, which is an outlier internationally in the push towards stratification (Usher 2016), it remains the case that the "top" universities (which have their own independent lobby group), receive a disproportionate share of research resources when measured, for example, against the percentage of students (including Doctoral students) they educate (U15 Group of Canadian Research Universities/Regroupement des universités de recherche du Canada 2016). In the much larger U.S. post secondary system, ten



universities received nearly 20% of all government research funds; as Weigley and Hess note, while these universities are among the richest in the country in terms of their endowments, public funding still constitutes the largest part of their R&D funding:

In addition to federal funding, the schools on this list also tend to have large endowments that support prestigious research facilities. All but one university on this list has an endowment of at least \$1 billion, and five of them had among the 15 largest endowments as of 2012. Stanford University's endowment of more than \$17 billion is the fourth largest in the United States, while the University of Michigan's is nearly \$7.7 billion and is the seventh highest.

While the schools on this list rely on massive endowments, the federal government comprised the majority of funding for R&D in all cases. At John Hopkins, 88% of the research budget came from federal funds. At the University of Pennsylvania, 80% of all R&D money came from federal funds. (2013)

Many have questioned the value of such an inequitable distribution of funds when a less concentrated, or less unequal, distribution could achieve greater outcomes. Dorothy Bishop argues, with respect to the UK's Research Excellence Framework (the REF), that there should be less of a disparity between rewarding research that is perceived to be "the best" and that which is perceived as merely average. Instead, Bishop argues, all research submitted to the REF should receive *some* funding and the perceived best research should receive proportionately less (Bishop 2013). This would have the benefit of decreasing the funding gulf between elite and middle-tier universities and would encourage diversity in the process. Of course such an approach may be politically troublesome for the academy. If funding is allocated on a scattered basis, following the logic that predictive approaches to quality are weak at best, then the authority claims of the university are substantially devalued.

There is, however, evidence to support the value of greater redistribution of research funding. Cook et al. showed that for UK Bioscience groups an optimal allocation of fixed resources would involve spreading the money between a larger number of smaller groups (Cook, Grange, and Eyre-Walker 2015). This was the case whether number of publications or number of citations were used as the measure of productivity. A similar conclusion is reached by Fortin and Currie who argue that scientific impact is only 'weakly money-limited' and that a more productive strategy would be to distribute funds based on 'diversity' rather than perceptions of excellence (Fortin and Currie 2013). Gordon and Poulin argued that, for science funding in Canada through the National Science and Engineering Research Council (NSERC, the main STEM funding agency), it would have cost less at a whole system level to simply distribute the average award to all

eligible applicants than to incur the costs associated with preparing, reviewing and selecting proposals (2009; although see Roorda 2009 for a critique of their calculation). A rough calculation of the system costs of preparing failed grant applications would suggest that they are in the same order of magnitude as research grant funding (Herbert et al. 2013)

We in fact can draw a spectrum of policy choices concerning the resourcing of research: from concentrating resources on the most deserving, allegedly “excellent,” institutions and researchers, to distributing them amongst all those that meet some minimum criteria—or even some subset, by lottery (Health Research Council of New Zealand 2016; Fang, Bowen, and Casadevall 2016). In the context of scarce resources and a desire to maximise outcomes there is even an argument for focusing most attention on the *worst* institutions; those that might most need resources in order to improve (Bishop 2013) and have the greatest scope for improvement. In this case, rather than excellence we would be looking for some sort of baseline level of qualification, “credibility” (Morgan 2016), perhaps, or “soundness.”

The problem with redistributive schemes is how to engage with politics. While proposing interesting and valuable thought experiments, they do not address the needs of working with governments who need to account for the distribution of public funds and may fear the optics of a purely distributive system. The narrative and the need for “excellence” and “international competitiveness” are important as a shared language of externally recognizable symbols that justify funding to government and to wider publics. Although as we showed in the first section, rhetorics of “excellence” are quite different in the many places that they are deployed, they are nonetheless used in a way which *assumes* a common meaning. Across a wide range of settings, “excellence” acts as a shared word, if not a truly shared concept, that bridges the gaps between different disciplines and stakeholder groups.

As we noted in the second section, this serves the interests of those who have already “earned” the label. The local construction of “excellence” is inherently conservative and maintaining its structures serves the interests of those who hold local power. Therefore, narratives arguing for redistribution need to be more than just interesting ideas and more than simply factually correct. They need to be compelling.

### Soundness over excellence

This is where a rhetoric built around “soundness” offers opportunities. This rhetoric could build on existing concerns about reproducibility and reporting. The idea that “sound research is good research” and that our focus should be on appropriate standards of description, evidence, and probity rather than flashy claims of superiority certainly has currency and can be developed. Such a narrative would also address deeper concerns regarding a breakdown in the culture of

research through its practice-driven audit culture and hypercompetition. “Soundness” resonates with public and funder concerns for value and it aligns with the need for improved communications and wider engagement encouraged by many governments and agencies.

### Successes in soundness

As we have just argued, a rhetoric focussed around soundness, on the idea that the most important quality of research is that it be properly done, can find a number of resonances with existing narratives. Concerns about reproducibility in the natural and human sciences have been a constant issue within the research community for the past several years as well as receiving wide exposure in the mainstream media (Open Science Collaboration 2015; Rehman 2013; Yong 2012b; Burman, Reed, and Alm 2010; Chang and Li 2015; Goldacre 2011; Lehrer 2010). While debates about the scale of the underlying problem continue (alongside the practicalities of publishing reproducibility studies), the importance of reproducibility as an underpinning quality of research is rarely questioned (for an unusual exception see Bissell 2013).

Reporting guidelines for animal experiments (Kilkenny et al. 2010) and clinical trials (Schulz et al. 2010) have similarly been an area of substantial discussion. Here the issue is of clarity of description but the underlying narrative of “doing things properly” is consistent. Work on registered replication studies in social psychology combines these concerns, providing a consistent process for describing the intent and process for attempting to replicate a previously published study, while guaranteeing publication of that result as a means of ensuring credit and recognition for contributing to this underpinning work (Simons, Holcombe, and Spellman 2014).

But the most notable success story of a narrative of “soundness” is the online journal, PLOS ONE and the set of journals that now follow its approach. PLOS ONE was launched with the stated aim of publishing any scientific research that was deemed technically sound, regardless of its perceived novelty or impact. This approach was made possible by two developments in academic publishing—the move to fully online publications without the need for print editions, and the growing acceptance of Article Processing Charge (APC)-funded Open Access as a viable publication model. These enabled the journal to consider and publish any manuscript that met its criteria, with no limitations on page space or fixed subscription revenue. As a result, the journal grew very quickly, becoming the largest journal in the world within five years of launching (MacCallum 2011).

The PLOS ONE model has been widely emulated, with almost every major scientific publisher now offering a journal with similar editorial criteria. This has created a competitive landscape with interesting properties. Traditional journals compete by seeking to publish the most “excellent” papers that they can attract. Those that succeed attract those papers that their

authors consider most important. Over time, success in this venture, its own form of hypercompetition, leads to a differentiated set of ranked journals driven by their own performative targets. PLOS ONE and its competitors also compete, but on sometimes different terms and in ways that arguably improve rather than imperil the research enterprise. Speed of publication, for example, always features in author surveys, and journals like PLOS ONE often advertise their average turnaround times. However, like traditional journals, they also compete on the basis of journal prestige, reputation and Impact Factor (Solomon 2014), albeit with an emphasis on soundness and lack of scarcity that reduces the risk of a “winner’s curse.” Even when the criteria for inclusion is only soundness, membership in the club of authors still provides a prestige benefit: that the doors of the club are open to all does not necessarily mean that there is no benefit to membership (Potts et al. 2016) .

This raises questions about the ability of researchers to see beyond “excellence.” Even when offered a distributive narrative of soundness, researchers often still find it difficult to avoid a concentrating rhetoric of “excellence.” The performance of “excellence,” the signalling of relative superiority through an additional line on the CV, is still more important from a career perspective than the science itself: nobody gets tenure for publishing to arXiv, no matter how good the quality of their research. And while reader attention or online conversation are gaining some currency as indicators of qualities valued in an article, the current discourse indicates that authors need to feel that they have cleared a higher bar than they in fact have. Many anecdotes from PLOS ONE authors involve being surprised by how tough the refereeing process was for their articles—a response that signals relative “excellence” that might otherwise not be apparent to the reader (see especially the comments to Curry 2012). A common complaint from the managers of journals such as PLOS ONE, indeed, is that their journals’ referees, who are usually made up of previous authors, often seek to reject papers that they feel do not meet their own perceptions of excellence, instead of focussing on the journal’s formal criterion of “soundness.” In other words, initiative like PLoS will have truly succeeded in changing researchers’ own bias towards (ultimately undemonstrable) “excellence” only when their rejection rate is seen to be less important than the evidence that controls are in place to ensure and encourage the recognition of “soundness.”

There are also questions to be raised about what any such criteria of “soundness” means in the context of the humanities. Martin Eve has suggested that this might work in a humanities paper that could “evince an argument; make reference to the appropriate range of extant scholarly literature; be written in good, standard prose of an appropriate register that demonstrates a coherence of form and content; show a good awareness of the field within which it was situated; pre-empt criticisms of its own methodology or argument; and be logically consistent” (Eve 2014, 144). More recently, Dan Morgan has suggested that “credibility” may be the humanities

equivalent of “soundness” (Morgan 2016). Yet, questions remain: are there shared conceptions of what properly done research looks like in the humanities? Should there be? How much does this differ across disciplinary boundaries, or even within disciplines where competing models and approaches can have profound differences in what counts as admissible evidence or as appropriate analysis or even of ethical practice.

Nonetheless the definition of research as a set of shared practices that have *standards* and that are carried out within the context of specific disciplinary cultures that support and enforce those standards provides a fertile ground for narratives built around soundness. The wider context of concerns about what is funded and making the case for responsible use of public funding spreads beyond reproducibility in the natural sciences to a clearer articulation of what constitutes useful or sound humanistic or social research.

Soundness therefore appears to be a plausible basis on which to build a new narrative, or rather to combine existing threads into a more consistent rhetorical framework. Such a framework will work to refocus our attention on research that is sufficiently valuable to be worth pursuing. To drive adoption and practice towards making this real, however, we will need more than a narrative. We will need resources to be redistributed towards supporting a broader class of research activities.

## Closing the loop: Planning for cultural change

We have advanced an argument that “excellence” is not just unhelpful to realising the goals of research and research communities but actively pernicious. A narrative of scarcity combined with “excellence” as a boundary object leads to concentration and thence hypercompetition. Hypercompetition in turn leads to performance of perceived “excellence,” driving a circular conservatism and reification of those performative qualities that are recognised as appropriate by existing power structures.

We have also argued that, while traditional jeremiads lay the blame for this at the feet of external actors—institutional administrators captured by neo-liberal ideologies, funders over-focussed on delivering measurable returns rather than positive change, governments obsessed with economic growth at the cost of social or community value—the roots of the problem in fact lie in our *internal* narratives and the nature of “excellence” and “quality” as boundary objects that *we* as researchers have developed into banners under which we march to protect our autonomy. The solution to such problems lies not in arguing for more resources for distribution via existing channels as this will simply lead to further concentration and hypercompetition. Instead, we have argued, we need to untangle our own narratives.

Finally, we have argued for a more pluralistic approach to the distribution of resources and credit. Where competition does take place it should do so on the basis of the many different *qualities*, plural, that are important to different communities using and creating research. But it should also be recognised that competition is not, in this context, an unalloyed good. In the context of assessing the risks of application of research Stirling and others argue for “broadening out *and* opening up” the technology assessment process (Ely, Van Zwanenberg, and Stirling 2014), that is to say increasing both the set of criteria considered and the range of people who have a voice in its assessment and application. The same approach needs to be applied to research assessment.

This leads to our argument for a focus on redistribution instead of concentration, which, we suggest, is necessary for three core reasons. Firstly because “excellence” cannot be recognised or defined consensually, except as a Wittgensteinian “beetle in a box” that no-one has ever seen, and even then, unlike Wittgenstein’s beetle-owners, by researchers who cannot agree even within disciplinary communities on which aspects of “excellence” might matter or be useful. Secondly we have argued for redistribution on its own merits; because concentration, and the hypercompetition to which it leads, break down our standards and cultures in systematic, predictable, and negative ways. And thirdly, we have shown that top-loading of research funding based upon anti-foundational principles of “excellence” is likely to hurt the incremental advances upon which research implicitly relies.

The argument for redistribution is a challenging one to advance. The rhetorics of scarcity, of concentration and competition are linked to strong cultural and economic narratives, particularly in the United Kingdom and United States. But as a route towards this goal we argue that we can build on existing narratives of “soundness” and “credibility”—which is to say on narratives of reproducibility, transparency, and high-quality reporting—in order to build a case for strong cultural practices that focus on fundamental standards that define proper scholarly and scientific practice. In taking this approach we root the discourse in long-standing traditions and culture, while also engaging with the newer concerns that align. It is through showing that we *can* recognise sound and credible research and that we *can* build strong cultures and communities around that recognition, that we lay the groundwork for making the case for redistribution. And *that* would be an excellent thing.

## Works cited

Alberts, Bruce, Marc W. Kirschner, Shirley Tilghman, and Harold Varmus. 2014. “Rescuing US Biomedical Research from Its Systemic Flaws.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (16): 5773–77.

- Aldhous, Peter. 2016. "Journal Rejects Studies Contradicting Precognition." *New Scientist*. Accessed February 19.  
<https://www.newscientist.com/article/dn20447-journal-rejects-studies-contradicting-precognition/>.
- Allington, Daniel, Sarah Brouillete, and David Golumbia. 2016. "Neoliberal Tools (and Archives): A Political History of Digital Humanities." *Los Angeles Review of Books*, May 1.  
<https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>.
- Alpher, R. A., H. Bethe, and G. Gamow. 1948. "The Origin of Chemical Elements." *Physical Review* 73 (7): 803–4.
- Anderson, Melissa S., Emily A. Ronning, Raymond De Vries, and Brian C. Martinson. 2007. "The Perverse Effects of Competition on Scientists' Work and Relationships." *Science and Engineering Ethics* 13 (4): 437–61.
- Andrade, Rodrigo de Oliveira. 2016. "Sharp Rise in Scientific Paper Retractions." *University World News*, January 8. <http://www.universityworldnews.com/article.php?story=20160108194308816>.
- Atkin, Philip A. 2002. "A Paradigm Shift in the Medical Literature." *BMJ* 325 (7378): 1450–51.
- Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. Fall 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences." *The Rand Journal of Economics* 42 (3): 527–54.
- Babbage, Charles. 1830. *Reflections on the Decline of Science in England: And on Some of Its Causes, by Charles Babbage (1830). To Which Is Added On the Alleged Decline of Science in England, by a Foreigner (Gerard Moll) with a Foreword by Michael Faraday (1831)*. B. Fellowes.
- Bauerlein, Mark. 2014. "Theory and the Humanities, Once More." *Inside Higher Ed*, November 13.  
<https://www.insidehighered.com/views/2014/11/13/essay-critiques-role-theory-humanities>.
- Belluz, Julia. 2016. "Do 'Top' Journals Attract 'Too Good to Be True' Results?" *Vox*. January 11.  
<http://www.vox.com/2016/1/11/10749636/science-journals-fraud-retractions>.
- Bem, D. 2011. "Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect." *Journal of Personality and Social Psychology* 100 (3): 407–25.
- Bertamini, Marco, and Marcus R. Munafò. 2012. "Bite-Size Science and Its Undesired Side Effects." *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 7 (1): 67–71.
- Birney, Ewan. 2012. "The Making of ENCODE: Lessons for Big-Data Projects." *Nature* 489 (7414): 49–51.
- . 2016. "Response on ENCODE Reaction." *Ewan's Blog: Bioinformatician at Large*. Accessed May 22. <http://genomeinformatician.blogspot.com/2012/09/response-on-encode-reaction.html>.
- Bishop, Dorothy. 2013. "The Matthew Effect and REF2014." *BishopBlog*. October 15.  
<http://deevybee.blogspot.ca/2013/10/the-matthew-effect-and-ref2014.html>.
- Bissell, Mina. 2013. "Reproducibility: The Risks of the Replication Drive." *Nature* 503 (7476): 333–34.
- Bollen, Johan, David Crandall, Damion Junk, Ying Ding, and Katy Börner. 2014. "From Funding Agencies to Scientific Agency." *EMBO Reports* 15 (2): 131–33.
- Boyle, Rebecca. 2013. "The Drama over Project Encode, and Why Big Science and Small Science Are Different." *Popular Science*, February 25.  
<http://www.popsci.com/science/article/2013-02/fracas-over-project-encode-and-why-big-science-and-small-science-are-different>.
- Brembs, Björn. 2015. "The Cost of the Rejection-Resubmission Cycle." *The Winnower*. doi:10.15200/winn.142497.72083.
- Brembs, Björn, Katherine Button, and Marcus Munafò. 2013. "Deep Impact: Unintended Consequences of Journal Rank." *Frontiers in Human Neuroscience* 7: 291.
- Brown, Titus. 22 May, 2016. "Titus Brown on Twitter." *Twitter*.  
<https://twitter.com/ctitusbrown/status/734393744640741379>.
- Burman, L. E., W. R. Reed, and J. Alm. 2010. "A Call for Replication Studies." *Public Finance Review* 38 (6): 787–93.
- Calcagno, V., E. Demoinet, K. Gollner, L. Guidi, D. Ruths, and C. de Mazancourt. 2012. "Flows of Research Manuscripts among Scientific Journals Reveal Hidden Submission Patterns." *Science* 338 (6110): 1065–69.

- Campanario, Juan Miguel. 1993. "Consolation for the Scientist: Sometimes It Is Hard to Publish Papers That Are Later Highly Cited." *Social Studies of Science* 23: 342–62.
- . 1995. "Commentary on Influential Books and Journal Articles Initially Rejected because of Negative Referees' Evaluations." *Science Communication* 16 (3): 304–25.
- . 1996. "Have Referees Rejected Some of the Most-Cited Articles of All Times?" *Journal of the American Society for Information Science (1986-1998)* 47 (4): 302–10.
- . 2009. "Rejecting and Resisting Nobel Class Discoveries: Accounts by Nobel Laureates." *Scientometrics* 81 (2): 549–65.
- Campanario, Juan Miguel, and E. Acedo. 2007. "Rejecting Highly Cited Papers: The Views of Scientists Who Encounter Resistance to Their Discoveries from Other Scientists." *Journal of the American Society for Information Science and Technology* 58: 734–43.
- Casadevall, Arturo, and Ferric C. Fang. 2012. "Winner Takes All." *Scientific American* 307 (2): 13.
- Chang, Andrew C., and Phillip Li. 2015. "Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say 'Usually Not.'" 2015-083. Finance and Economics Discussion Series. Washington: Board of Governors of the Federal Reserve System.  
<http://www.federalreserve.gov/econresdata/feds/2015/files/2015083pap.pdf>.
- Chen, Chaomei, Zhigang Hu, Jared Milbank, and Timothy Schultz. 2013. "A Visual Analytic Study of Retracted Articles in Scientific Literature." *Journal of the American Society for Information Science and Technology* 64 (2): 234–53.
- Chubb, Jennifer, and Richard Watermeyer. 2016. "Artifice or Integrity in the Marketization of Research Impact? Investigating the Moral Economy of (pathways To) Impact Statements within Research Funding Proposals in the UK and Australia." *Studies in Higher Education* 0 (0): 1–13.
- Cicchetti, Domenic V. 1991. "The Reliability of Peer Review for Manuscript and Grant Submissions: A Cross-Disciplinary Investigation." *The Behavioral and Brain Sciences* 14 (01): 119–35.
- Claxton, Larry D. 2005. "Scientific Authorship. Part 1. A Window into Scientific Fraud?" *Mutation Research* 589 (1): 17–30.
- Cook, Isabelle, Sam Grange, and Adam Eyre-Walker. 2015. "Research Groups: How Big Should They Be?" *PeerJ* 3 (June): e989.
- Costello, Leslie C. 2010. "Perspective: Is NIH Funding the 'Best Science by the Best Scientists'? A Critique of the NIH R01 Research Grant Review Policies." *Academic Medicine: Journal of the Association of American Medical Colleges* 85 (5): 775–79.
- Curry, Stephen. 2012. "PLoS ONE: From the Public Library of Sloppiness?" *Reciprocal Space*. April 1.  
<http://occamstypewriter.org/scurry/2012/04/01/plos1-public-library-of-sloppiness/>.
- Dean, Dwight G. 1989. "Structural Constraints and the Publications Dilemma: A Review and Some Proposals." *The American Sociologist* 20 (2): 181–87.
- De Man, Paul. 1982. "The Resistance to Theory." *Yale French Studies*, no. 63: 3–20.
- de Solla Price, D. J. 1975. *Science Since Babylon*. Yale University Press.
- Dickersin, Kay. 2005. "Publication Bias: Recognizing the Problem, Understanding Its Origins and Scope, and Preventing Harm." In *Publication Bias in Meta-Analysis*, 9–33. John Wiley & Sons, Ltd.
- Dickersin, Kay, S. Chan, T. C. Chalmers, H. S. Sacks, and H. Smith. 1987. "Publication Bias and Clinical Trials." *Controlled Clinical Trials* 8 (4): 343–53.
- Dobbs, David. 2006. "Trial and Error." *The New York Times*, January 15.  
[http://www.nytimes.com/2006/01/15/magazine/15wwln\\_idealab.html?\\_r=0](http://www.nytimes.com/2006/01/15/magazine/15wwln_idealab.html?_r=0).
- Edwards, Richard. 2015. "Bioinformatics Is Just like Bench Science and Should Be Treated as Such." *The Cabbages of Doom*. August 21.  
<http://cabbagesofdoom.blogspot.com/2015/08/bioinformatics-is-just-like-bench.html>.
- Ely, Adrian, Patrick Van Zwanenberg, and Andrew Stirling. 2014. "Broadening out and Opening up Technology Assessment: Approaches to Enhance International Development, Co-Ordination and Democratisation." *Research Policy* 43 (3): 505–18.
- Eve, Martin Paul. 2014. *Open Access and the Humanities : Contexts, Controversies and the Future*. Cambridge: Cambridge University Press.
- "Excellence in Research for Australia." 2016. *Australian Research Council*. Accessed March 26.  
<http://www.arc.gov.au/excellence-research-australia>.
- Eyre-Walker, Adam, and Nina Stoletzki. 2013. "The Assessment of Science: The Relative Merits of



- Post-Publication Review, the Impact Factor, and the Number of Citations." Edited by Jonathan A. Eisen. *PLoS Biology* 11 (10): e1001675.
- Fanelli, Daniele. 2009. "How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data." *PloS One* 4 (5): e5738.
- . 2011. "Negative Results Are Disappearing from Most Disciplines and Countries." *Scientometrics* 90 (3): 891–904.
- Fang, Ferric C., Anthony Bowen, and Arturo Casadevall. 2016. "NIH Peer Review Percentile Scores Are Poorly Predictive of Grant Productivity." *eLife* 5 (February). doi:10.7554/eLife.13323.
- Fang, Ferric C., R. Grant Steen, and Arturo Casadevall. 2012. "Misconduct Accounts for the Majority of Retracted Scientific Publications." *Proceedings of the National Academy of Sciences of the United States of America* 109 (42): 17028–33.
- Fitzpatrick, Kathleen. 2011. *Planned Obsolescence*. New York: New York University Press.
- Flaherty, Colleen. 2014. "Doing Themselves In?" *Inside Higher Ed*, October 20.  
<https://www.insidehighered.com/news/2014/10/20/conference-speakers-say-liberal-arts-must-return-purer-form-survive>.
- Fortin, Jean-Michel, and David J. Currie. 2013. "Big Science vs. Little Science: How Scientific Impact Scales with Funding." Edited by Vincent Larivière. *PloS One* 8 (6): e65263.
- Gans, Joshua S., and George B. Shepherd. 1994. "How Are the Mighty Fallen: Rejected Classic Articles by Leading Economists." *The Journal of Economic Perspectives: A Journal of the American Economic Association* 8 (1): 165.
- Goldacre, Ben. 2011. "I Foresee That Nobody Will Do Anything about This Problem." *Bad Science*. April 23.  
<http://www.badsience.net/2011/04/i-foresee-that-nobody-will-do-anything-about-this-problem/>.
- Gonon, François, Jan-Pieter Kongsman, David Cohen, and Thomas Boraud. 2012. "Why Most Biomedical Findings Echoed by Newspapers Turn out to Be False: The Case of Attention Deficit Hyperactivity Disorder." *PloS One* 7 (9): e44275.
- Goodrich, Dorris West. 1945. "An Analysis of Manuscripts Received by the Editors of the American Sociological Review from May 1, 1944 to September 1, 1945." *American Sociological Review* 10 (6): 716–25.
- Gordon, Richard, and Bryan J. Poulin. 2009. "Cost of the NSERC Science Grant Peer Review System Exceeds the Cost of Giving Every Qualified Researcher a Baseline Grant." *Accountability in Research* 16 (1): 13–40.
- Graur, Dan, Yichen Zheng, Nicholas Price, Ricardo B. R. Azevedo, Rebecca A. Zufall, and Eran Elhaik. 2013. "On the Immortality of Television Sets: 'Function' in the Human Genome According to the Evolution-Free Gospel of ENCODE." *Genome Biology and Evolution* 5 (3): 578–90.
- Grieneisen, Michael L., and Minghua Zhang. 2012. "A Comprehensive Survey of Retracted Articles from the Scholarly Literature." *PloS One* 7 (10): e44118.
- Groppe, David M. 2015. "Combating the Scientific Decline Effect with Confidence (intervals)." *BioRxiv*. Cold Spring Harbor Laboratory. doi:10.1101/034074.
- Hassell, M. P., and R. M. May. 1974. "Aggregation of Predators and Insect Parasites and Its Effect on Stability." *The Journal of Animal Ecology* 43 (2): 567–94.
- Health Research Council of New Zealand. 2016. "Explorer Grants." *Health Research Council*. Accessed February 19.  
<http://www.hrc.govt.nz/funding-opportunities/researcher-initiated-proposals/explorer-grants>.
- Herbert, Danielle L., Adrian G. Barnett, Philip Clarke, and Nicholas Graves. 2013. "On the Time Spent Preparing Grant Proposals: An Observational Study of Australian Researchers." *BMJ Open* 3 (5): e002800.
- Herrnstein Smith, Barbara. 1988. *Contingencies of Value: Alternative Perspectives for Critical Theory*. Harvard University Press.
- Higher Education Funding Council for England. 2016. "Funds for Research." Higher Education Funding Council for England. <http://www.hefce.ac.uk/funding/annalocns/1516/research/>.
- Hill, Helene, and Joel Pitt. 2014. "Failure to Replicate: A Sign of Scientific Misconduct?" *Publications* 2 (3). Multidisciplinary Digital Publishing Institute: 71–82.

- Hoover, W. G., B. Moran, B. L. Holian, H. A. Posch, and S. Bestiale. 1988. "Computer Simulation of Nonequilibrium Processes." In *Shock Waves in Condensed Matter 1987*, edited by S. C. Schmidt and N. C. Homes, 191–94. North-Holland.
- Hoover, W. G., H. A. Posch, and S. Bestiale. 1987. "Dense-fluid Lyapunov Spectra via Constrained Molecular Dynamics." *The Journal of Chemical Physics* 87 (11): 6665–70.
- Ioannidis, John P. A., Kevin W. Boyack, and Richard Klavans. 2014. "Estimates of the Continuously Publishing Core in the Scientific Workforce." *PloS One* 9 (7): e101698.
- Ioannidis, John P. A., Kevin W. Boyack, Henry Small, Aaron A. Sorensen, and Richard Klavans. 2014. "Bibliometrics: Is Your Most Cited Work Your Best?" *Nature* 514 (7524): 561–62.
- Jay, Paul. 2014. "How Not to Defend the Liberal Arts." *Inside Higher Ed*, October 27. <https://www.insidehighered.com/views/2014/10/27/essay-state-liberal-arts>.
- Kennedy, Donald. 2004. "The Old File-Drawer Problem." *Science* 305 (5683): 451.
- Kilkenny, Carol, William J. Browne, Innes C. Cuthill, Michael Emerson, and Douglas G. Altman. 2010. "Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research." *PLoS Biology* 8 (6): e1000412.
- King, Molly, Shelley Correll, Jennifer Jacquet, Carl T. Bergstrom, and Jevin D. West. 2015. "Men Set Their Own Cites High: Gender and Self-Citation across Fields and over Time." Palo Alto, CA: Working paper. Retrieved from <http://www.eigenfactor.org/gender/self-citation/SelfCitation.pdf>. <http://www.eigenfactor.org/gender/self-citation/>.
- King, Molly, Jevin D. West, Jennifer Jacquet, Shelley Correll, and Carl T. Bergstrom. 2014. "Gender Composition of Scholarly Publications." *Eigenfactor*. <http://www.eigenfactor.org/gender/self-citation/>.
- Kohn, Alfie. 1999. *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Boston: Houghton Mifflin Co.
- Kuhn, Thomas S. 1963. "The Essential Tension: Tradition and Innovation in Scientific Research." In *Scientific Creativity: Its Recognition and Development*, edited by Calvin Walker Taylor and Frank Barron, 341–54. Robert E. Krieger Publishing Company.
- Kuhn, Thomas S., and Ian Hacking. 2012. *The Structure of Scientific Revolutions*. Fourth edition. Chicago ; London: The University of Chicago Press.
- Labbé, Cyril. 2010. "Ike Antkare One of the Great Stars in the Scientific Firmament." *ISSI Newsletter* 6 (2): 48–52.
- Lamont, Michèle. 2009. *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, Mass: Harvard University Press.
- Larsen, Peder Olesen, and Markus von Ins. 2010. "The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index." *Scientometrics* 84 (3): 575–603.
- Lawrence, Peter A. 2007. "The Mismeasurement of Science." *Current Biology* 17 (15): R583–85.
- Lehrer, Jonah. 2010. "The Truth Wears Off." *New Yorker*, December 13. <http://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off>.
- Lerner, Josh, and Julie Wulf. 2006. "Innovation and Incentives: Evidence from Corporate R&D." w11944. National Bureau of Economic Research. <http://www.nber.org/papers/w11944.pdf>.
- Lindner, Mark D., and Richard K. Nakamura. 2015. "Examining the Predictive Validity of NIH Peer Review Scores." Edited by Neil R. Smalheiser. *PloS One* 10 (6): e0126938.
- Lord, Robert G., Christy L. de Vader, and George M. Alliger. 1986. "A Meta-Analysis of the Relation between Personality Traits and Leadership Perceptions: An Application of Validity Generalization Procedures." *The Journal of Applied Psychology* 71 (3): 402.
- Lowe, Derek. 2013. "Farewell to Bioinformatics." *In the Pipeline*. January 30. [http://blogs.sciencemag.org/pipeline/archives/2013/01/30/farewell\\_to\\_bioinformatics](http://blogs.sciencemag.org/pipeline/archives/2013/01/30/farewell_to_bioinformatics).
- Lundh, Andreas, Marija Barbateskovic, Asbjørn Hróbjartsson, and Peter C. Gøtzsche. 2010. "Conflicts of Interest at Medical Journals: The Influence of Industry-Supported Randomised Trials on Journal Impact Factors and Revenue - Cohort Study." *PLoS Medicine* 7 (10): e1000354.
- MacCallum, Catriona J. 2011. "Why ONE Is More Than 5." *PLoS Biology* 9 (12): e1001235.
- Marche, Stephen. 2012. "Literature Is Not Data: Against Digital Humanities." *Los Angeles Review of Books*, October 28. <https://lareviewofbooks.org/article/literature-is-not-data-against-digital-humanities/>.

- Marcovitch, Harvey. 2010. "Editors, Publishers, Impact Factors, and Reprint Income." *PLoS Medicine* 7 (10): e1000355.
- Matzinger, P., and G. Mirkwood. 1978. "In a Fully H-2 Incompatible Chimera, T Cells of Donor Origin Can Respond to Minor Histocompatibility Antigens in Association with Either Donor or Host H-2 Type." *The Journal of Experimental Medicine* 148 (1): 84–92.
- Meng, Weishi. 2016. "Peer Review: Is NIH Rewarding Talent?" *Science Transparency*. January 10. <https://sciencetransparency.wordpress.com/2016/01/10/peer-review-is-nih-rewarding-talent/>.
- Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159 (3810): 56–63.
- . 1973. "The Normative Structure of Science." In *The Sociology of Science: Theoretical and Empirical Investigations*, edited by Norman W. Storer, 267–78. Chicago and London: University Of Chicago Press.
- . 1988. "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property." *Isis; an International Review Devoted to the History of Science and Its Cultural Influences* 79 (4): 606–23.
- Miller, J. Hillis. 1987. "Presidential Address 1986. The Triumph of Theory, the Resistance to Reading, and the Question of the Material Base." *Publications of the Modern Language Association* 102 (3): 281–91.
- Miller, Stephen. 2015. *Excellence and Equity: The National Endowment for the Humanities*. University Press of Kentucky.
- Moran, Bill, William G. Hoover, and Stronzo Bestiale. 2016. "Diffusion in a Periodic Lorentz Gas." *Journal of Statistical Physics* 48 (3-4): 709–26. Accessed May 22.
- Morgan, Dan. 2016. "Lessons Learned, and How the Landscape Has Already Changed." presented at the Open Access @ UNT/Library Publishing Forum, University of North Texas, May 19. <https://openaccess.unt.edu/symposium/2016/live-streaming-oa-uc-press-lessons-learned-and-how-landscape-has-already-changed>.
- Multiple contributors. 2013. "A Farewell to Bioinformatics (2012)." *Hacker News*. January 30. <https://news.ycombinator.com/item?id=5123022>.
- Munafò, M. R., G. Stothart, and J. Flint. 2009. "Bias in Genetic Association Studies and Impact Factor." *Molecular Psychiatry* 14 (2): 119–20.
- Nature. 2013. "Form and Function." *Nature News* 495 (7440): 141.
- Networks of Centres of Excellence of Canada, Communications. 2015. "Networks of Centres of Excellence." April 13. [http://www.nce-rce.gc.ca/index\\_eng.asp](http://www.nce-rce.gc.ca/index_eng.asp).
- Neylon, Cameron. 2010. "It's a Discovery Deficit: Sydney University 19 November 2010." *Science in the Open*. November 19. <http://cameronneylon.net/presentations/its-a-discovery-deficit-sydney-university-19-november-2010/>.
- Nicholson, Joshua M., and John P. A. Ioannidis. 2012. "Research Grants: Conform and Be Funded." *Nature* 492 (7427): 34–36.
- Nielsen, Michael A. 2012. *Reinventing Discovery: The New Era of Networked Science*. Princeton, N.J.: Princeton University Press.
- Nosek, Brian A., Jeffrey R. Spies, and Matt Motyl. 2012. "Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability." *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 7 (6): 615–31.
- Nowotny, Helga. 2014. "Excellence Attracts Excellence--and What about the Rest? Reflections on Excellence and Inclusion." presented at the EMBO-EMBL Anniversary Science and Policy Meeting, Heidelberg, July. [http://www.helga-nowotny.eu/downloads/helga\\_nowotny\\_b160.pdf](http://www.helga-nowotny.eu/downloads/helga_nowotny_b160.pdf).
- Nurse, Paul. 2015. "Ensuring a Successful UK Research Endeavour." BIS/15/625. London: Department for Business, Innovation & Skills. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/478125/BIS-15-625-ensuring-a-successful-UK-research-endeavour.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/478125/BIS-15-625-ensuring-a-successful-UK-research-endeavour.pdf).
- O'Connor, Pat, and Clare O'Hagan. 2015. "Excellence in University Academic Staff Evaluation: A Problematic Reality?" *Studies in Higher Education*, February, 1–15.
- O'Donnell, Daniel Paul. 2012. "Editors' Choice: 'There's No Next about It': Stanley Fish, William Pannapacker, and the Digital Humanities as Paradiscipline Dpod Blog." *Digital Humanities Now*,

September.

<http://digitalhumanitiesnow.org/2012/09/editors-choice-theres-no-next-about-it-stanley-fish-william-pannapacker-and-the-digital-humanities-as-paradiscipline-dpod-blog/>.

———. 2015. "Could We Design Comparative Metrics That Would Favour the Humanities?" *Daniel Paul O'Donnell*. March 29.

<http://people.uleth.ca/~daniel.odonnell/Teaching/could-we-design-comparative-metrics-that-would-favour-the-humanities>.

OECD. 2014. "Chapter 6. The German Excellence Initiative." In *Promoting Research Excellence*, 145–63. OECD Publishing.

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349 (6251): aac4716.

Pagano, Michele. 2006. "American Idol and NIH Grant Review." *Cell* 126 (4): 637–38.

Pannapacker, William. 2011a. "Pannapacker at MLA: Digital Humanities Triumphant?" *Brainstorm*. January 8.

<http://chronicle.com/blogs/brainstorm/pannapacker-at-mla-digital-humanities-triumphant/30915>.

———. 2011b. "'Big Tent Digital Humanities,' a View From the Edge, Part 1." *The Chronicle of Higher Education*, July 31. <http://chronicle.com/article/Big-Tent-Digital-Humanities/128434/>.

———. 2012. "'No DH, No Interview.'" *The Chronicle of Higher Education*, July 22.

<http://chronicle.com/article/No-DH-No-Interview/132959/>.

Peters, Douglas P., and Stephen J. Ceci. 1982. "Peer-Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again." *The Behavioral and Brain Sciences* 5 (02): 187–95.

Potts, Jason, John Hartley, Lucy Montgomery, Cameron Neylon, and Ellie Rennie. 2016. "A Journal Is a Club: A New Economic Model for Scholarly Publishing." *Social Science Research Network*, April. doi:10.2139/ssrn.2763975.

Psych Filedrawer. 2016. "The Filedrawer Problem." *PsychFileDrawer.org*. Accessed February 19. <http://www.psychfiledrawer.org/TheFiledrawerProblem.php>.

Readings, Bill. 1996. *The University in Ruins*. Harvard University Press.

Rehman, Jalees. 2013. "Cancer Research in Crisis: Are the Drugs We Count on Based on Bad Science?" *Salon*, September 1. [http://www.salon.com/2013/09/01/is\\_cancer\\_research\\_facing\\_a\\_crisis/](http://www.salon.com/2013/09/01/is_cancer_research_facing_a_crisis/).

Resnik, David B., and Gregg E. Dinse. 2013. "Scientific Retractions and Corrections Related to Misconduct Findings." *Journal of Medical Ethics* 39 (1): 46–50.

Resnik, David B., Elizabeth Wager, and Grace E. Kissling. 2015-7. "Retraction Policies of Top Scientific Journals Ranked by Impact Factor." *Journal of the Medical Library Association: JMLA* 103 (3): 136–39.

Roberts, Leslie. 2001. "Controversial From the Start." *Science* 291 (5507): 1182–88.

Roderick, G. K., and R. G. Gillespie. 1998. "Speciation and Phylogeography of Hawaiian Terrestrial Arthropods." *Molecular Ecology* 7 (4): 519–31.

Rogan, Peter K., and Guang Yong Zou. 2013. "Best Practices for Evaluating Mutation Prediction Methods." *Human Mutation* 34 (11): 1581–82.

Roorda, Sjoerd. 2009. "The Real Cost of the NSERC Peer Review Is Less than 5% of a Proposed Baseline Grant." *Accountability in Research* 16 (4): 229–31.

Rothstein, Hannah R. 2014. "Publication Bias." In *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd.

Schooler, Jonathan. 2011. "Unpublished Results Hide the Decline Effect." *Nature* 470 (7335): 437.

Schulz, Kenneth F., Douglas G. Altman, David Moher, and CONSORT Group. 2010. "CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials." *BMJ* 340 (March): c332.

Science, Public Library of. 2016. "Who We Are." *PLoS*. Accessed May 12. <https://www.plos.org/who-we-are>.

Secord, J. A. 2015. *Visions of Science: Books and Readers at the Dawn of the Victorian Age*. University of Chicago Press.

Şekercioğlu, Çağan H. 2013. "Citation Opportunity Cost of the High Impact Factor Obsession." *Current Biology* 23 (17): R701–2.

- Sewitz, Sven. 2014. "The Excellence Agenda Is a Trojan Horse for Austerity." *Research*. September 3. [http://www.researchresearch.com/index.php?option=com\\_news&template=rr\\_2col&view=article&articleId=1346207](http://www.researchresearch.com/index.php?option=com_news&template=rr_2col&view=article&articleId=1346207).
- Shirky, Clay. 2010. *Cognitive Surplus: Creativity and Generosity in a Connected Age*. New York: Penguin Press.
- Siler, Kyle, Kirby Lee, and Lisa Bero. 2015. "Measuring the Effectiveness of Scientific Gatekeeping." *Proceedings of the National Academy of Sciences* 112 (2): 360–65.
- Simons, Daniel J., Alex O. Holcombe, and Barbara A. Spellman. 2014. "An Introduction to Registered Replication Reports at Perspectives on Psychological Science." *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 9 (5): 552–55.
- Smith, Richard. 2006. "Research Misconduct: The Poisoning of the Well." *Journal of the Royal Society of Medicine* 99 (5): 232–37.
- Solomon, David J. 2014. "A Survey of Authors Publishing in Four Megajournals." *PeerJ* 2 (April): e365.
- Star, Susan Leigh. 1989. "The Structure of Ill-Structured Solutions: Boundary Objects and Heterogeneous Distributed Problem Solving." In *Distributed Artificial Intelligence*, edited by Michael N. Huhns and Les Gasser, 37–54. 2. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- . 2010. "This Is Not a Boundary Object: Reflections on the Origin of a Concept." *Science, Technology & Human Values* 35 (5): 601–17.
- Steen, R. Grant. 2011. "Retractions in the Scientific Literature: Is the Incidence of Research Fraud Increasing?" *Journal of Medical Ethics* 37 (4): 249–53.
- Stephan, Paula. 2014. *How Economics Shapes Science*. [S.l.]: Harvard Univ Press.
- Sterling, Theodore D. 1959. "Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance—Or Vice Versa." *Journal of the American Statistical Association* 54 (285): 30–34.
- Stilgoe, Jack. 2014. "Against Excellence." *The Guardian*, December 19. <https://www.theguardian.com/science/political-science/2014/dec/19/against-excellence>.
- Stirling, Andy. 2007. "A General Framework for Analysing Diversity in Science, Technology and Society." *Journal of the Royal Society, Interface / the Royal Society* 4 (15): 707–19.
- "Sustaining Excellence Awards." 2016. *Wellcome Trust*. Accessed May 19. <http://www.wellcome.ac.uk/Funding/Public-engagement/Funding-schemes/Sustaining-Excellence-Awards/index.htm>.
- Tartamelia, Vito. 2014. "The True Story of Stronzo Bestiale (and Other Scientific Jokes)." *Parolacce*. October 5. <http://www.parolacce.org/2014/10/05/the-true-story-of-stronzo-bestiale/>.
- The Economist*. 2015. "The World Is Going to University." <http://www.economist.com/news/leaders/21647285-more-and-more-money-being-spent-higher-education-too-little-known-about-whether-it>.
- The University of Tokyo Global Center of Excellence. 2016. "Global Center of Excellence for Mechanical Systems Innovation." *The University of Tokyo Global COE*. Accessed May 12. <http://www.u-tokyo.ac.jp/coe/english/list/category2/base7/summary.html>.
- Tian, Xuan, and Tracy Yue Wang. 2011. "Tolerance for Failure and Corporate Innovation." *The Review of Financial Studies*, December. doi:10.1093/rfs/hhr130.
- Tijdsink, Joeri K., Reinout Verbeke, and Yvo M. Smulders. 2014. "Publication Pressure and Scientific Misconduct in Medical Scientists." *Journal of Empirical Research on Human Research Ethics: JERHRE* 9 (5): 64–71.
- U15 Group of Canadian Research Universities/Regroupement des universités de recherche du Canada. 2016. "Our Impact." *U15*. Accessed May 18. <http://u15.ca/our-impact>.
- University Grants Commission. 2016. "Universities (UPE)." *University Grants Commission*. Accessed February 19. [http://www.ugc.ac.in/page/Universities-\(UPE\).aspx](http://www.ugc.ac.in/page/Universities-(UPE).aspx).
- University of Arizona Commission on the Status of Women. 2015. "Avoiding Gender Bias in Reference Writing." Tucson AZ: University of Arizona. [http://www.csw.arizona.edu/sites/default/files/csw\\_2015-10-20\\_lorbias\\_pdf\\_0.pdf](http://www.csw.arizona.edu/sites/default/files/csw_2015-10-20_lorbias_pdf_0.pdf).
- "USC Center for Excellence in Teaching." 2016. Accessed May 12. <http://cet.usc.edu/>.
- Usher, Alex. 2016. "Massification Causes Stratification." *Higher Education Strategy Associates*. May 5.

- <http://higherstrategy.com/massification-causes-stratification/>.
- Vinkers, Christiaan H., Joeri K. Tjldink, and Willem M. Otte. 2015. "Use of Positive and Negative Words in Scientific PubMed Abstracts between 1974 and 2014: Retrospective Analysis." *BMJ* 351 (December): h6467.
- Wang, Jian, Reinhilde Veugelers, and Paula E. Stephan. 2016. "Bias Against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators." *Social Science Research Network*, January. doi:10.2139/ssrn.2710572.
- Wankel, C. 2013. *Digital Humanities: Current Perspective, Practices, and Research*. Cutting-Edge Technologies in Higher Education. Emerald Group Publishing Limited.
- Weigley, Samuel, and Alexander E. M. Hess. 2013. "Universities Getting the Most Government Money." *247wallst.com*. April 25.  
<http://247wallst.com/special-report/2013/04/25/universities-getting-the-most-government-money/>.
- Weller, Ann C. 2001. *Editorial Peer Review: Its Strengths and Weaknesses*. Information Today.
- Wikipedia contributors. 2016. "Literary Theory." *Wikipedia, The Free Encyclopedia*. May 17.  
[https://en.wikipedia.org/w/index.php?title=Literary\\_theory&oldid=720648584](https://en.wikipedia.org/w/index.php?title=Literary_theory&oldid=720648584).
- Wilson, Andrew. 2011. "Failing to Replicate Bem's Ability to Get Published in a Major Journal." *Notes from Two Scientific Psychologists*. May 7.  
<http://psychsciencenotes.blogspot.ca/2011/05/failing-to-replicate-bems-ability-to.html>.
- Wittgenstein, Ludwig. 2001. *Philosophical Investigations: The German Text, with a Revised English Translation*. Blackwell.
- World Bank. 2016. "Research and Development Expenditure (% of GDP)." *The World Bank | Data*. Accessed May 25. <http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS>.
- Yong, Ed. 2012a. "A Failed Replication Draws a Scathing Personal Attack from a Psychology Professor." *Not Exactly Rocket Science*. March 10.  
<http://blogs.discovermagazine.com/notrocketscience/2012/03/10/failed-replication-bargh-psychology-study-doyen/#.VsZpH0Leezc>.
- . 2012b. "Replication Studies: Bad Copy." *Nature* 485 (7398): 298–300.
- Young, Neal S., John P. A. Ioannidis, and Omar Al-Ubaydli. 2008. "Why Current Publication Practices May Distort Science." *PLoS Medicine* 5 (10). doi:10.1371/journal.pmed.0050201.
- Young, S. Stanley, and Heejun Bang. 2004. "The File-Drawer Problem, Revisited." *Science* 306 (5699): 1133–34; author reply 1133–34.
- Zankl, Heinrich. 2003. *Fälscher, Schwindler, Scharlatane*. Erlebnis Wissenschaft. Wiley-VCH Verlag.