Towards an analytical understanding of peer review in research funding

Thomas König

Acknowledgements

Many scattered thoughts have been noted in my diary between 2010 and 2013, when I was observing panel meetings at the ERC. A stay at Harvard’s Center for European Studies over the summer of 2014 allowed me to gain overview of the peer review literature. My assessment of the odd world of peer review greatly benefited from discussions with Stefan Beljean, Svend-Erik Skaaning, Lea Haller, Ulrike Felt, Christian Fleck, and Michèle Lamont. An internal discussion on peer review and transparency at the FWF in December 2014, and a presentation of an earlier draft of this paper at the University of Vienna’s Department of Science and Technology Studies KISS seminar series in February 2015 have had similar impact on shaping this paper. It could not have been written, however, without the generous support of the Austrian-American Fulbright Commission and the Swedish Riksbankens Jubileumsfond.
1. Introduction

In his monograph on “real science”, John Ziman has called peer review a “key institution of scientific culture” (Ziman 2000, 246). This assumption is probably widely shared today: according to numerous accounts, peer review is the “gold standard”, and the equivalent for science to what representative democracy is to modern societies. Other than his fellow scientists, however, Ziman was critical of applying peer review to the fringes of the scientific culture. To him, it was a means to maintain quality assurance within that culture (such as, deciding as to whether to publish an article in a journal or not). Applying this principle to the allocation of public funds for research would mean that it were used for making policy predictions about future research, something for which Ziman did not find it suited at all (Ziman 2000, 185–189; Ziman 2002).

Despite of Ziman’s high-minded protest, peer review in research funding has long become the most revered way of decision-making for allocating funds to research. As the allocation of (public) funds is often based on peer review, it also affects the daily routines of researchers across the globe. Today, they spend a good share of their professional work either for writing proposals to be reviewed, or for reviewing their peers’ proposals. The impact on science as a whole is often thought to be similarly profound, which is why there is nowadays a considerable interest in peer review in research funding. As we will see in the next section, many studies on different aspects of peer review have been conducted. At the same time, many critics have summoned the near end of peer review, either because of overburdening the system, or because of failing accuracy (Nicholson and Ioannidis 2012).

However, critics and analysts have both failed to come up with an argument in order to understand the success of peer review in the first place. Success, I hasten to add, does not mean that peer review is key in making the right funding decisions, but the fact that peer review is commonly regarded as doing so. The following chapter examines the growing body of literature on peer review, showing that no adequate conceptualization for analysing peer review in research funding exists so far. Chapter 3 then proceeds a) to define what peer review is; b) at which particular point in time it entered the realm of research funding, and c) how it functions. It argues, first, that albeit being the same decision-making principle, peer review in research funding plays a significantly different role than in other domains (most particularly, in journal
editing); second, that the implementation of peer review for allocating research funds was a historical coincidence; and third, that the great advantage of peer review is its procedural flexibility. Building on these three legs, chapter 4 makes the core theoretical argument that, besides its many symbolic meanings, which it has acquired over time, the ubiquity of peer review relies on its power to legitimize funding decisions in parallel to scientists and to policy-makers. It is its ability to create dual legitimacy that cannot be matched by any other decision-making principle, and hence makes peer review such an admired tool for allocating funds to research. The final chapter then discusses some prospects for further research on the topic.

2. Research on peer review in research funding

Scientists and policy-makers who wonder whether peer review is actually the best system of allocating public funds to scientific research have stimulated research on peer review in research funding. Under this paradigm (see also Table 1), we can roughly distinguish two approaches: One concerned with potential procedural limitations of the process, and the other concerned with the outcome of specific funding calls. We can briefly glance over the first approach (no. 1 in Table 1), where studies are close to the institutions conducting peer review in research funding (such as funding agencies); their main aim is to provide guidance for properly setting up peer review processes and to refine them in order to become better (cf. Dearfield and Flaak 2000). The second approach (no. 2 in Table 1), which is actually the most prominent one, takes a more “scientific” stance. It is mainly concerned with quantitatively assessing data from the results of peer reviewing and then making conclusions about reliability, validity, and fairness of the process. Not by incidence, the first major study in this realm was conducted because of nagging doubts about those issues (S. Cole, Rubin, and Cole 1978; J. R. Cole and Cole 1981). Studies in this (mainstream) tradition illuminate to which extent the overall outcome of peer review is without bias, relies on the reviewers’ judgment, and funds the right kind of people; thus, they aim at analysing the integrity of a specific peer review implementation at a given funding agency based on the post-ex
assessment. Some studies have made news and keep on being quoted, such as the famous one on the Swedish funding agency showing that the peer review process has been driven by “nepotism and sexism” (Wennerås and Wold 1997). Many other studies have used similar methodological approaches on different sets of data and have come either to similar or opposing results (Bornmann, Mutz, and Daniel 2007; Sandström and Hällsten 2007; Mutz, Bornmann, and Daniel 2012).

Table 1: Overview of different streams of analysing peer review in research funding

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>No.</th>
<th>Focus</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is peer review fair/efficient?</td>
<td>1</td>
<td>Instruments</td>
<td>Normative description of rules</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Procedures / Post-ex assessment of single calls</td>
<td>Descriptive and regressive statistics</td>
</tr>
<tr>
<td>Why is peer review accepted?</td>
<td>3</td>
<td>Procedures / Peer reviewers</td>
<td>Interviews, non-participatory observation, comparative</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Instruments / Peer review criteria</td>
<td>Assessing comparatively or in-depth</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Organisation</td>
<td>Single case, multiple methods</td>
</tr>
</tbody>
</table>

Compiled by the author.

The debate on whether ex ante peer review is fair, reliable, and valid, remains open until today. However, a recent meta study comes to the conclusion that “the retrospective design used in most of the available studies, the specificity of their settings and the questionable quality of the methods used, seriously limit the possibility of deriving general information even on these marginal aspects investigated.” (Demicheli 2008, 6). One reason (that is often mentioned) why this mainstream of studies lacks in depth concerns access to data.1 But even when data is available, studies are methodologically sophisticated, but theoretically poor: All they usually do is to measure the outcome of a given call against the backdrop of statistical artefacts, such as the intraclass correlation coefficient, control variables, etc. (Hirschauer 2004; Reinhart 2012, 98–122) Because they are only concerned with the (political) question whether peer review is fair, reliable, and based on the right values, those studies have a great value for proponents of peer review (such as funding

---

1 The secrecy of most funding agencies about data is notorious. More often than not, the organisations actually carrying out the peer review are either not able (due to heavy workload) or not willing (due to protecting their turf) or not allowed (due to legal provisions) to hand out data on various steps during the process.
agencies); but their contribution to receiving at a social scientific understanding of peer review in research funding is negligible.

Partly due to this dissatisfactory situation, social scientists have started to take a different look at peer review. Their focus of attention has shifted from whether peer review is fair/efficient to the question why peer review is generally accepted as the best mechanism in allocating research funding in the first place. This is a crucial step forward, as it turns away from an essentially political question (concerned with values such as good, fair, effective) to a social scientific one (concerned with assessing the legitimacy of how a specific mechanism accrues such values in the first place); instead of taking peer review as a given, it looks at it as a specific “social control institution” (Langfeldt 2006, 31) We can roughly distinguish three approaches of research under this new paradigm that all have emerged over the past decade or so.

One distinctive approach (no. 3 in Table 1) is mostly concerned with values meshed into the process. Studies here aim at revealing the (rather informal) criteria applied by the reviewers when judging proposals (cf. Lamont 2009; Huutoniemi 2012). Thus, the studies provide insights into the “intersubjective understandings” of quality (Lamont and Huutoniemi 2011) of academic scholars and researchers when asked to take over the role as “peers”. Besides the fact that they have a great strategic value to future applicants who want to know what their evaluators may find important in their proposals, those studies also break down what is usually expected to be a scientifically professional attitude during the peer review process. While the findings are reassuring to some extent, they also reveal a good portion of negotiation and tactical manoeuvring on behalf of the peers.

Another approach (no. 4 in Table 1) is about the criteria and political implications and to which extent these can actually influence the results of different funding streams. An early version has been concerned with how “organizational constraints influence review outcome of peer review” (Langfeldt 2001, 822; also Langfeldt 2006). In a comparison of a new generation of funding streams that attempt to incorporate political goals such funding “high risk” and “unconventional” research into their peer review procedure, the study finds that “the decision process itself tends to be rather risk averse” (Heinze 2008, 316). Similarly, when a group of U.S.-based scholars inquired whether other criteria than intrinsically scientific ones can be made use of in grant peer review in order to achieve lofty political goals, such as societal impact (Frodeman and Parker 2009; Holbrook and Frodeman 2011), their findings are
ambiguous: on the one hand, mainstreaming wider social criteria into the peer review process is often arbitrary, but on the other hand, bringing in those criteria does have a “dedisciplining” of peer review (Frodeman and Briggle 2012; Holbrook 2010b).

Two things can be learnt from the studies assessed so far. First, the studies focusing on reviewers (no. 3) remind us that peer review is a mechanism in which scientifically intrinsic values play an important role, while the studies on funding calls (no. 4) find that these values are nonetheless embedded in a larger scheme (typically, a specific funding stream) and that this scheme, as research policy instrument, carries political intentions. In other words: funding streams are set up with specific political intentions, but if the allocation of funding is based on peer review, it relies heavily on scientific reasoning nonetheless. The two approaches provide insights into different aspects of various components of the process, but they fail to erect a systematic framework of how those different components work together, and thus do not come up with a comprehensive theoretical argument as to why peer review is ubiquitous.

Second, and on a more general level, it becomes clear that peer review – taken with the full ambiguity of the notion – is itself a place where the negotiation between science and policy is taking place. Thus, it is important to distinguish between two layers of empirical evidence: in most of the cases, studies concentrate on data deriving from actual calls. This layer, then is about the actual evidence produced by a given peer review procedure. Analysing this layer provides interesting insights into the practices within peer review, but provides little evidence about the negotiation process between the two worlds. That process can only be found by focusing on the second layer, namely the various components and their definition of peer review procedures. This layer, I argue, constitutes itself a separate body of evidence, that is, the policy layer. If we want to understand the success of peer review, it is necessary to theoretically conceptualize this second layer.

Our literature review is not yet completed. Only recently, a case study on the Swiss National Science Foundation (SNF) has made a serious attempt to look at the entire process of peer review in research funding in order to answer this question (Reinhart 2012).² The findings of this study (no. 5) are remarkable, not only because they elegantly reject the methodological obsession of quantitative assessments which we

---

² Reinhart (2012, 193) states in the final remarks of his study: „Die Irritation, die das Forschungsinteresse dieser Arbeit hauptsächlich begründet, liegt darin, dass Peer Review ein universeller Tatbestand der modernen Wissenschaft zu sein scheint."

6
briefly discussed before, but also because they provide a sociological account of peer review as a procedure in research funding, a procedure that consists of a partly open, partly secretive consecutive row of decision-making steps. At each of those steps, specific types of knowledge are brought in, based on which different actors can make informed decisions. As a whole, the procedure is organized in a way that it boils down from panoply of different knowledge items to a binding, binary decision. (Reinhart 2012, 131) Eventually, Reinhart claims that it is the stability of peer review that explains its overall success (2012, 153). With stability, he means that peer review produces a reliable result, the list of projects to be funded.

3. Dual legitimacy

However, as ambitious as Reinhart’s study may be, the result is all but convincing. Partly, this is because of the fact that his analysis concentrates too much on the organisational framework (the SNF) that manages and conducts the peer review procedure. It is true that, for the organization that actually is built around this process and relies on it, stability is an essential feature. But peer review in research funding is not done for its own sake; stability is a necessary, but insufficient feature to explain the ubiquity of peer review in research funding. Policy-makers are not interested in stability per se; they want to be rest assured that public money is spent efficiently. And scientists are not necessarily interested in stability either; they want to be sure that they receive money and, if they do not receive it, that there are scientifically comprehensible reasons why they do not receive it.4 Reinhart’s study, like most of the other work that we have examined in the previous chapter, rests on a flawed concept: it does not provide a viable definition of what peer review in research funding actually is, that is: what is it deployed to achieve?

3 Only in passing – since this is not the objective of this contribution – it should be noted that Reinhart’s final assessment of his research object remains somewhat ambiguous, as also specifically noted by a reviewer (Simet 2013).
4 Reinhart himself seems to be aware of this crucial fact, since he later refers to the “very solid basis of legitimacy” [“eine äußerst solide Legitimationsbasis”] that peer review has been able to acquire; but he rather vaguely introduces the new notion of “meta-stability” [“Metastabilität”] to explain the fact that “peer review decisions are understood as knowledge-based and democratic at the same time” [“indem Peer Review-Entscheidungen gleichzeitig als wissensbasiert und demokratisch verstanden werden.”] (Reinhart 2012, 193).
3.1 The achievement of peer review in research funding

To answer this question, it is best to start with a brief historical examination of the deployment of peer review for the allocation of funds to research. With the exception of forerunners, most notably in the UK and in Germany, peer review was established on a large scale for the first time immediately after World War II, and not incidentally this took place when governments – with the U.S. leading the field – began pouring public funds into research. Ironically, the link between public funding for research and the rise of peer review in the allocation of those funds seems to be so obvious (Holbrook 2010a, 323–324),⁵ that it is seldom recognised. At that time, the general relationship between science and policy was characterized by what Dan Sarewitz later called the “leap of faith” (1996, 10); Guston summarized “the of the emerging policy field of science policy as follows:

“the political community agrees to provide resources to the scientific community and to allow the scientific community to retain its decision-making mechanisms and in return expects forthcoming but unspecified technological benefits.” (Guston 2000, 62)⁶

My argument here is that peer review almost archetypically translated this “dominant ideology” into practice (Van der Meulen 1998, 405). In his book on “Local Justice” (1992; see also 1995),⁷ political theorist Jon Elster is broadly concerned with the question of how scarce goods are allocated in a way that is acceptable by those immediately involved. His case studies involve college admission, organ transplants, etc. Elster develops a useful catalogue of terminologies. Most importantly, he distinguishes between a general “principle” from the “procedures” that have been implemented “as the operational version of a principle” (Elster 1992, 62–63).⁸ We can conclude that one reason why there is so much confusion about defining peer review seems to be that, usually, the term is used to characterize the principle and the procedure at the same time. Indeed, this is the problem of the studies in chapter 2.

⁵ Elsewhere, I have described this as a new “funding regime” (König 2015).
⁶ This is also sometimes called the „social contract for science“.
⁷ For a comprehensive overview of justice theories, see Konow (2003). A recent overview of research in the sociology of valuation and evaluation is provided by Michèle Lamont (2012). Both accounts give special credit to Elster’s attempt to come up with a “descriptive” (as Konow puts it), rather than a normative, account of different types of principles and mechanisms to allocate scarce goods.
⁸ In passing, it should be noted that, while peer review as a principle is used in many different instances, such as journal editing or clinical assessments, only in research funding it is used for allocating scarce goods.
As a principle, peer review describes a distinctive way of decision-making, namely, that a decision has to be based on the assessment (review) of experts (the peers) of a certain artefact (a research paper) or practice (someone’s teaching skills). In most abstract terms, the principle of peer review is always the same: it aims at quality assurance and improving credibility without taking other (worldlier) values into account. However, the principle can be deployed to accomplish different tasks, and depending on this task, it takes on different meaning. Most notably, when it comes to research funding, the principle is no longer situated within the “scientific culture” (to borrow Ziman’s phrase again, even though it is normatively laden), but at the fringes of that culture and the realm of policymaking. That has serious consequences on what the principle has to achieve.

3.2 Principle and procedure of peer review

Today, the most common deployment of peer review within the scientific culture concerns journal editing. Scientific journals rely on peer review in order to decide whether a submitted draft article should be published or rejected. The achievement of peer review is to legitimise that decision. It rests entirely on scientific terms, which is of course the great strength of the principle in the first place. In research funding, the principle is the same, but its achievement is different. Again, the decision is ultimately based on scientific terms; but, because it is tasked to legitimise the allocation of funds, funds that come from outside the scientific culture, it promises to achieve dual legitimacy – both to the world of politics and to the scientific culture.

Table 2: Different deployments of peer review principle

<table>
<thead>
<tr>
<th>Peer review</th>
<th>… in journal editing</th>
<th>… in research funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deployed</td>
<td>Within scientific culture</td>
<td>Border btw science &amp; politics</td>
</tr>
<tr>
<td>Meaning</td>
<td>Maintain quality of journal</td>
<td>Allocate (public) funds</td>
</tr>
<tr>
<td>(Promised) achievement</td>
<td>Single legitimacy</td>
<td>Dual legitimacy</td>
</tr>
</tbody>
</table>

Compiled by the author.

Dual legitimacy of peer review means that, by deciding on the allocation of funds to research, it has to live up to different expectations of scientists and policy-makers,
respectively: the scientific world is primarily interested in the fairness of the decision-making, while the world of policy-making is concerned with its efficiency (Gurwitz, Milanesi, and König 2014). Fairness means that the funding decision is based on scientific judgment of proposals that have all been formally treated in the same way. Efficiency is a more complex matter; it can actually mean three things: that the allotment itself raises the respect of policy-makers among scientists (who, after all, are key stakeholder in the field of science policy); that the funding is allotted in a cost-effective way; and that the funded research yields results that justify the investment. We will return to the issue of efficiency later; for the moment, it is important to understand that, in order to achieve them, it requires an actual procedure. The principle of peer review presents the prospect of legitimacy, but it is the procedure that brings it in.

In a famous study in the late 1960s, German sociologist Niklas Luhmann tried to understand why, in a modern democratic system and under the rule of (positivist) law, procedures had taken over such a prominent role. Its title – “legitimation through procedure” (Luhmann 1969) – makes his general argument clear. Peer review, once implemented, becomes a procedure in the same sense, and hence mimics the general logic as determined by Luhmann. Its purpose is “to convert indeterminacy into determinacy”;

“a procedure must lead to some kind of outcome (a decision, the constitution of something, or knowledge). In between the first and the last task, processes of negotiation, selection, and verification have to take place, step by step.” (Lorenz 2009, 8).

From this point of view, peer review as a procedure is “programmed”, meaning that the allocation of funds is to be based “on the already existing foundations for decision-making” (Lorenz 2009, 6). That foundation is the principle of peer review – that decisions be taken based on scientific (and not worldlier) values.

3.3 The meaning of the peer review procedure

---

9 In passing, it should be noted that this is exactly what Reinhart found in his study of peer review (without referring to Luhmann).
10 „Programming procedures“, on the other hand, are far more complex, because they „must first create such foundations and therefore must first determine who is to decide (political election) or what is to be decided upon (legislation).” (Lorenz 2009, 6)
We now see how principle and procedure depend on each other – and why scholars often have difficulties in distinguishing between the two. But how is it possible for peer review – as a procedure – to combine (again in Elster’s terminology) “substantive properties of individuals”, or “criteria”, and “discretionary decision-making”, that is, a mechanism that gives intermediaries (peers, in our case) “a great deal of latitude in interpreting the general guidelines they are to implement” (Elster 1992, 62–63) in a way that fairness and efficiency are indeed achieved (and that means: that scientists and politicians accept it to be fair and efficient)? Continuing the comparison of different situations of peer review, it is possible to look at differences between the procedures of peer review for journal editing and for allocating funds (Table 3). The comparison yields some interesting differences between the two modes of procedures, which further amplify our general conjecture that peer review within the scientific culture is distinctively different from a deployment of the principle on the fringes of that culture.

In journal editing, the implementation of peer review requires that research described in the submitted text be already done when the review is conducted. As a procedural methodology, peer review then assesses each individual submission on its own. The value of scientific quality is defined mostly by the originality and the (methodical) rigidity of the submitted paper. The (rather simple, at least in comparison) procedures of peer review in journal editing rely on different modes of shopping for reviews – often the procedure is called a “double blind peer review”, meaning that reviewers and authors are kept anonymous to each other – and the rigor with which the reviews are then applied.

Table 3: Differences between two modes of peer review procedures

<table>
<thead>
<tr>
<th>Peer review …</th>
<th>… in journal editing</th>
<th>… in research funding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Object(s)</strong></td>
<td>Finalized scientific work</td>
<td>Proposals of future works; CV</td>
</tr>
<tr>
<td><strong>Methodology</strong></td>
<td>Single case</td>
<td>Comparison of proposals</td>
</tr>
<tr>
<td><strong>Scientific quality</strong></td>
<td>Method, originality</td>
<td>Feasibility, promise</td>
</tr>
<tr>
<td><strong>Result</strong></td>
<td>Simplicity</td>
<td>Complexity</td>
</tr>
</tbody>
</table>

Compiled by the author.

The procedure of peer review in research funding rests on a distinctively different prerequisite. Since it determines the allocation of externally appropriated funds to a selection of proposed scientific projects, the review assesses outlines for future
(research) activities, and typically involves the CV of the applicant (as well as other objects). The value of scientific quality in the assessment focuses on the (scientific) promises of the proposed project and also its feasibility. In order to allocate the available funding to what are deemed the best proposals, the procedure rests on a relational methodology: it pools and compares assessments of several proposals. Because of the need to achieve fairness and efficiency at the same time, the procedure relies on at least six components that constitute a list of minimum requirements for any peer review in research funding. It is important to understand that those components all are necessary requirements to achieve dual legitimacy. Fairness requires a clearly defined process, that is, a consecution of steps starting with who is eligible to apply for funding and with what kind of proposal, up to making sure that every submission will be treated equally during the review; similarly, it needs to be clear that the peers are impartial. In order to achieve efficiency, the procedure is usually outsourced to a separate organizational framework; and the criteria for reviewing have to be revealed.

4. Orchestration of peer review in research funding

So far, we have found that, in its initial deployment, peer review congenially fit the “dominant ideology” of science policy (Guston 2000). Two aspects are important here. Once peer review had entered research policy, it flourished. As other industrialized countries quickly followed the new hegemon (Krige 2006), spending public funds for research became a trend across the Western hemisphere, as can be seen from the fact that “intermediary organizations” (Braun 1997) emerged in almost all those countries.¹¹ As it is typical for policy instruments, those intermediaries – or organizational frameworks for the implementation of peer review, as we have called them in the previous chapter – developed their own interest in staying alive. That’s why most studies on science policy have been greatly stimulated by principal-agent theory (Braun and Guston 2003), which helped to explain how this policy field could

¹¹ There have been notable exceptions, such as Austria (cf. Stampfer, Pichler, and Hofer 2010).
establish a certain degree of stability despite the fact that “the relation between government and science is basically unstable” (Van der Meulen 1998, 399).

The second important issue is that peer review turned into something like a mantra for scientists. It was emboldened by programmatic statements, such as Michael Polanyi’s famous assumption that,

“[s]o long as each allocation [of funds, TK] follows the guidance of scientific opinion, by giving preference to the most promising scientists and subjects, the distribution of grants will automatically yield the maximum advantage for the advancement of science as a whole.” (Polanyi 1962, 61)

In other words, peer review was quickly conferred a status similar to a set of other abstract concepts that emerged with the discourse on science policy, such as “basic research” (Schauz 2014), or the “value-free ideal for science” (Douglas 2009, 44–65).

For scientists, it became a matter of course to allocate funds with peer review as the decision-making principle.

Even as science policy developed further, and the period of “blind delegation” was replaced by other modes of policy-making, first the “delegation by incentives” in the 1960s, and “austerity”, “contracts”, and “delegation to networks” in the decades afterwards (Braun 2003), peer review remained in place. With the tides of science policy, focus diverted from project-based funding to other forms of funding. Partly, other allocations principles, such as research institution-building (block grants), person-centred funding (prizes), and secretive distribution (aka clientilism) were (re-)established; partly, peer review was integrated into a wider set of principles for allocating funds to predefined research topics or other conditions (such as transnational collaboration). Many attempts to replace or emulate the principle failed, and peer review retained a (more or less) strong foothold.

As already noted, the meaning of “efficiency” is necessarily ambiguous and consists of at least three different meanings – one being the self-interest of the politicians in charge, the next being focused on cost-efficiency, and the third expecting results of the funded research. As a decision-making principle that rests on scientific judgment, peer review can probably achieve the two first meanings; for the third, however, it can only make promises at best. Thus, the achievement of efficiency can always be put in doubt. Nonetheless, peer review was successful in its niche funding basic (academic) research. The reason for that is not only because of its institutional residues and of the ideological surplus that it had created. Even though it is unable itself to fully achieve all three meanings of efficiency, overall it is still better suited to achieve dual
legitimacy than other allocation principles that have been developed so far. The main reason is that its claim of fairness cannot be met by any other procedure. As much as politicians have been striving for a tighter grip on defining the terms of allocating the funds to research in the past, the residual argument – that only scientists can identify projects that deserve funding – cannot be counteracted.

4.1 A new paradigm in science policy

Peer review quickly became the most revered principle of allocating funds. Over the time, however, different shifts in research funding impacted on peer review’s overall prominence, as science policy developed into a complex set of different “frames” (Ulnicane 2015) and “conceptual frameworks” (Godin 2009), with finding new means to stimulate “knowledge production” (Hessels and van Lente 2008). From the 1990s on, research and development (R&D, as the common abbreviation goes) also became a key priority for policy-making in many regions of the world (cf. Laredo and Mustar 2001). Besides other innovations such as “New Public Management” entering universities (Clark 2004), competition amongst researchers and research institutions was deemed politically attractive as a way to stimulate the overall productivity of the “national innovation system”.

First and foremost, allotment of public resources to science increased significantly under the new paradigm, and the research landscape diversified. Competition among intermediary organizations increased to explain why their mode of allocating funds to research are different (and better) than that of others. Funding agencies that long had run their procedure now became more pro-active and strategic, with peer review as their unique selling point (Rip 1994). This was mostly an uncoordinated campaign, but it was immensely successful: Peer review was heralded as a stamp of excellence, and it was now also emulated by institutions that had not actively relied on the allocation principle before.

However, the political realm demanded more oversight as well as proof of investments yielding benefits not only in academic terms, but also to the greater good. And scientists, no longer securely embedded at their universities with life-long contracts but instead forced to acquire third party funding (either because they had to
make their living from it, or because they needed it to keep their position at their research institution), increasingly applied to peer reviewed calls because of the (academic) prestige that went along with it. For funders, thus, to continuously achieve dual legitimacy was becoming more intricate. As Arie Rip realized during the early phase of this new period,

“[R]esearch councils must legitimate next year’s budget in terms of an attractive portfolio of projects-to-be-funded, while the money they get must be disbursed in a way that is acceptable to the scientists.” (Rip 1994, 12)

As a consequence, refining the procedure has become even relevant, and funding agencies have further cultivated their specific implementations of peer review. It is now a complex “orchestration” (Rip 1994, 18) of different steps, preparatory sessions, rules and regulations.

This includes that, besides its initial meaning – to produce a clear result – the peer review procedure has acquired a second one, that is, to (actively and passively) mediate different interests and, in general terms, to satisfy the two equal “principals” (policy makers and scientists) (Braun and Guston 2003). In other words, the procedure is tasked to absorb interests, doubts and criticism, while the principle still lends the funding decision its specific character, namely, that the allocation is based on an assessment based on scientific values.

4.2 Gateways for negotiation

With the changing paradigm in the 1990s, ironically, peer review became ubiquitous because it occupied a more or less stable niche in the landscape of research funding. It continues to be successful because the procedural components of peer review in research funding are flexible enough to intercept attempts to interfere, and is capable to resolve doubts about the effectiveness and the fairness of the principle. But this is not enough to understand actually what is going on. The before-mentioned distinction in two meanings of the procedure (to produce a result vs. to mediate different interests) allows us to first understand better peer review as a procedure in general, before we learn how it manages to accomplish the second meaning as well.

As we have seen before, peer review remains an anachronism in the research funding landscape, as long as we do not acknowledge the capacity of the procedure to absorb
nagging doubts and inquiries from the principals. Furthermore, I have argued that we need to understand this capacity as a distinct second meaning of the procedure, one that had only grown when the principle and its implementations have long been established in research funding. More specifically, I suggest understanding the different components (briefly mentioned in the previous section) not only in regards to their operational means, but also as opportunities to mediate interests and to absorb doubts about the general decision-making principle. Looked at the peer review procedure from this point of view, it is easy to realize not only its flexibility and adaptability in terms of operationalization, but also that, by way of implementation, defining and refining each of the four facilitating components automatically offers gateways for mediation (Table 4).

![Table 4: Components, role, operational elements, and gateways for negotiations](image)

(1) Organisation: All funding streams are embedded in an organisational framework organisational setting, and those settings differ in types – Braun (1993), for example, distinguishes between “mission-agencies” and “all-round agencies” – and along the legislative framework in which they are implemented. Even if we do not bother about the fact that national legislation provides different rights and obligations to the organisational bodies enveloping each funding stream, we have to accept that the independence of those organisational bodies may differ, for example when it comes to the right to define their own funding program or to conduct the peer review procedure. Probably most importantly, that independence also varies regarding available budget. For example, the ERC has a (rather secured) budget for seven years in a row, while the NSF has to negotiate its budget with the US Congress every year
anew. In a nutshell, thus, structural details about the organisational framework as well as its degree of independence have to be taken into account.

(2) Process: The next component concerns the procedure put in place. Here, two elements of operationalization can be distinguished: The sequences into which the procedure is divided, and the mode of breaking up the streams into topics. As for sequences, we can identify procedures that follow several steps, sometimes involving different types of reviewers and the proposers themselves (with interviews or site visits); furthermore, some procedures also foresee the possibility of objecting, and re-assessing a proposal that has been turned down. As for the second element, since it is usually necessary to break streams down into fields in order to make sure that proposals from a given area are handled by people with the required expertise, we can find different modes of handling this. To stick to our comparison with NSF and ERC again, the former breaks down its allocation of funds already at organisational level, meaning that there are separate Directorates for different areas. The ERC, on the other hand, ensured topical division only at the rather soft level of the procedures themselves.

(3) Criteria: The next component concerns the criteria along which project proposals are to be assessed and evaluated. A critical operational element here concerns the purpose of those criteria: Are they directed at assessing only the intrinsic scientific value of the project and of its proposer(s), or are there also others asked for? For example, the ERC brags about asking for “excellence only”; the NSF, on the other hand, has a standard set of criteria, which involve societal impact as well. Another element concerns the objects that are to be assessed: both the NSF and the ERC follow a very narrow approach, that is, only the project proposal and the principal investigator’s; however, others may also ask for assessing the host institution as well (as was the case with the ERC until a few years ago).

(4) Peers: Finally, the last component concerns the peers who are hired to do the evaluations. Again, we can distinguish at least two operational elements: One is the policy of how peers are selected, which is often done on an ad hoc basis (NSF), but can also follow more or less rigid rules, involving formal criteria, such as gender balance, geographical spread, etc. (ERC). The other element concerns the role of the

---

12 The following remarks are only for illustration and far from exhaustive; the data for the two case studies, ERC and NSF, were selected for a paper I wrote with Michael Gorman, comparing those two funding agencies in reacting to the need to focus on interdisciplinary research (König and Gorman 2016).
peers; basically, two different ones can be distinguished: typically, there are two, and peers can be asked to act either as external reviewers (asked to assess one specific proposal), or as member of a specifically designed panel that is tasked to comparatively assess all assigned proposals. The ERC and the NSF actually use both roles, but other funding agencies follow a different strategy.

4.3 Mediating concerns

It is important to differentiate between two different strategic approaches of orchestrating peer review: one concerns safeguarding the procedural integrity of the allocation principle, the other concerns maintaining the legitimacy of the result of the procedure. As for the first, in times of increasing number of applications and pressure for winning competitive funding, funding agencies have to make sure that the peer review procedure is not overwhelmed by number of applications and is not threatening away reviewers due to the workload. There are different strategies of funding agencies to deal with the matter: they can either minimize the number of candidates eligible for submitting a proposal, or they can minimize the size of applications, or they can use several stages for eliminating applications. Safeguarding the integrity is mostly a matter of applying different techniques.

It is, however, the second approach that interests us here: mediating political concerns during the peer review. The notion of (political) mediation requires quick clarification: What is mediated here are interests formulated by different stakeholders, not necessarily only from the political realm, but also from the scientific community and from the organisation (the funding agency), the peers, and the governing body. The overall ambition of the mediation is to make sure that, by way of addressing a concern during the procedure, the dual legitimacy of the principle of peer review is not impaired. By way of refining the components and their operational elements, interests can be mediated and, therefore, absorb criticism and nagging doubts: at organisational level, refinement typically concerns different aspects of autonomy; in regards to the process, it concerns its overall purpose and the role of feedback loops; the formulation of criteria is often a political issue; and so are the selection policy and the decision-making power of the peers. This list is probably not exhaustive, but what
it aims to indicate is that mediations can be situated with all four components that are means of the procedure.

The gateways for mediation relate to four different concerns (or interests) that aim at influencing the peer review procedure, according to what they actually address: the outcome of the decision-making, which concerns all requests for formal equality and non-discrimination; the conduct of the procedure, which demands openness; the (promised) impact of the funded research, which asks for means to alter the way the scientific enterprise is conducted; and the values underlying the funded research, which asks to bring them in line with the values generally expressed by the public.

1. Probably the first concern regarding peer review in research funding also in historical terms was about the (formal) outcome of the peer review procedure. It had been a major concern at least since the allegations against the NSF in the 1970s; and since the study on a gender bias in Swedish research funding (mentioned in chapter 2), this topic is a classic. All funding agencies are obsessed with proving that their mode of allocating funds is fair and efficient to that end, which is of course a sign that there are many reasons to doubt this. Gender is one example; another concerns geographical distribution of funds or distribution along scientific domains.

2. The second concerns doubts about the process itself. Why is there so much secrecy in place? Transparency has become the promise of the day (Heald 2006). It is expected to solve the access and, in its most radical version, it wants to change the system of peer review altogether (Gurwitz, Milanesi, and König 2014; Mietchen 2014).

3. Then there are the doubts about the impact of peer review funded research. It has often been mentioned that peer review is primarily a conservative decision-making utility, since it relies mostly on disciplinary denominations and standards. The power of the notion of interdisciplinarity (Lyall et al. 2013) rests on the promise that, in one way or another, this sort of research is new, innovative, creative and original (Heinze 2008) and thus yields more (and better) results. Funding agencies have fully subscribed to fund interdisciplinary research, despite the problems of achieving it (König and Gorman 2016).

4. Finally, there is also the concern about the fact of compliance of values established in research proposals with what is believed to be societal values.
What is interesting is that each of the four approaches finds its advocates in the world of science as well as in the world of policy-making. It may thus seem as if the issue of mediating interests could eventually result in features contributing to an ever more robust procedure of peer reviewing. But on second thought, this perception does not take into account that the interests formulated in the mediation are different. For scientists, the obstacle is always fairness. For policy-makers, their main focus is on efficiency. For example, if the gender distribution of a result of a peer reviewed funding call is imbalanced, policy-makers are concerned primarily because it is a sign of not distributing efficiently (namely, to all those theoretically eligible), while for scientists, it is primarily because it means for them personally, it might be disadvantageous. The allegation is the same – that there is a gender bias – but the interest behind it is different.

It can be further broken down as related to the costs of the procedure and the contentment among scientists that is established. The efficiency in terms of the costs derives from the fact that much of the work that is involved with peer review is put on the shoulders of scientists. And, since scientists themselves are the core group addressed in science policy, their contentment is an additional benefit that must not be forgotten.

5. Conclusions

What, then, follows from this conceptual framework? First of all, for the past 20 years or so, peer review has been recognized as an essential mechanism in the continuous efforts of science and policy to be involved with each other (Rip and Van der Meulen 1996). It is time for studies concerned with analysing those efforts (and the resulting frictions) to take this into account, that is, to abolish the idea that peer review is something that is solely in the hands of scientists and to start investigating more in detail the various attempts to reign in and to counteract. On a very general level, peer review has always been a negotiation table for the two worlds involved. But with the changes in academia and the changing relationship between science and policy, it has

---

13 As mentioned in the previous chapter, only a few studies have recently started to look at instruments and compare them based on some rather randomly chosen criteria (Langfeldt 2006; Heinze 2008; Holbrook 2010b).
also fallen under increasing scrutiny. Thus, we need to understand where peer review is going. How much refinement of the implementations is possible? How much credibility does it buy?

There are also two very clear analytical consequences. One is directed at the hermeneutical exercise of interpreting implementations of peer review. If we understand peer review as the place for negotiating science and policy-making, the way it is implemented also provides information about the more subtle intentions of specific research funding streams and the usages of research funding; since those two are (like a simplified model of “input” and “outtake”) in constant tension, we also have to understand that herein lies one of the fields where the negotiations between science and policy-making takes place.

The other consequence is more in terms of modelling and concerns the question for the threshold at which peer review will be regarded as too costly by either of the involved principals. This relates to the fact that the specific benefits of peer review have a counterweight. To scientists, the major burden is the time that has to be invested not only for writing applications, but also for reviewing them. There are also costs to the public, however: politicians have to delegate funding allocation to scientists, and they have to deal with potential fall-out of the uneven distribution of funds. But, as in the case of the scientists, peer review’s specific benefit to the public again weighs more than those costs. That benefit can be summarized as its effectiveness for allocating funds to research, and it is based on three factors: a high probability of identifying the “right” proposals; concealment of costs of conducting peer review in the scientific world; and high credibility of the principle among scientists (see Table 5).

Table 5: Benefits and costs of peer review in research funding

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Scientific world</th>
<th>Political world</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairness: formal scientific criteria trump</td>
<td>Effectiveness: Concealment of costs; identifying the “right” proposals; credibility gain</td>
<td></td>
</tr>
<tr>
<td>Time investment</td>
<td>Delegation of decision-making power</td>
<td></td>
</tr>
</tbody>
</table>

Compiled by the author.

Even if, on a general level, the benefits outweigh the costs, it is possible to arrive at a situation where the specific costs are perceived to be higher than the than the specific
benefits – either because the costs are growing, or because the benefits are declining. For a long time already, doubts about the procedural fairness of the peer review procedure flare up (for a quick history of those doubts, cf. Ziman 1981; Van der Meulen 1998). More recently is the phenomenon that the increase of writing and reviewing proposals is said to reach a critical level threatening to overwhelm the scientific system as a whole. (Herbert et al. 2013) Similarly, there have been doubts about the efficiency of identifying the ‘right’ proposals. (Nicholson and Ioannidis 2012) And policy makers have started to ask more bluntly for evidence of the societal impact of peer-review funded research (Holbrook and Frodeman 2011). However, and despite many calls for an end of peer review, the mechanism is still in place, and, to some extent, more firmly so than ever. The reason why both sides continue to estimate the benefits to be higher is that, albeit the examples just given seem to indicate the opposite, recent developments in science policy have put peer review center stage.

The fact that peer review implementations are not only flexible in regards to their institutional environment, but also in regards to the challenges presented in the previous chapter, should lead us to acknowledge that peer review is uniquely suited as the primary allocation principle for research funding. It is, nonetheless, interesting to ask to which extent the reconfigurations of peer review implementations due to the various challenges have actually also changed the conduct of scientific practices and their organisation. The final question, thus, refers us back to the initial starting point of research on grant peer review: Does it do what it promises? Or does it simply fund those proposals that are the best or the cleverest written or that are submitted by the old-boys network? We will never know.
Bibliography


