This is a pre-copyedited, author-produced version of an article accepted for publication in Research Evaluation following peer review. The version of record Axel Philipps, Dealing with potentials and drawbacks of peer review panels: About the intertwined layers of determinacy and indeterminacy, *Research Evaluation*, 2024 is available online at: https://doi.org/10.1093/reseval/rvae024.

Dealing with potentials and drawbacks of peer review panels: About the intertwined layers of determinacy and indeterminacy

Axel Philipps

Abstract

To improve evaluation processes in science, scholars of grant peer review and science policy often problematize that factors such as emotions, group dynamics, and informal arrangements influence panel discussions. They emphasize their negative effects on scientifically grounded deliberations and search for solutions to decrease such impacts. By doing this, these scholars easily play down positive effects of emotion work and informal talks during panel sessions and seem less encouraged to study the interplay of formally organized and more informal exchanges. To take both into consideration, in this essay I outline a concept of review processes as intertwined layers of determinacy and indeterminacy. It proposes that determinate outcomes of panel discussions (e.g., definite judgements, funding recommendations) can only be understood relative to the process's indeterminacy (contextual vagueness such as informal talk, emotion work, tacit compromises). This theoretical framework (a) will help analytically to investigate indeterminacy in review panel processes as a whole as well as in panelists' situated scientific reasonings and (b) will generate knowledge for more effective evaluation management.

Keywords

Peer review, panels, evaluation, informal talk, emotion, mitigating behavior

Peer review is a constitutive part of science and is pivotal for its legitimacy. It is the principal mechanism of quality control in science through the assessment of scientific ideas and findings by fellow scientists. Most stakeholders in science not only believe in its effectiveness but it is also assumed that only scientists who do research in the same field of inquiry are able to produce scientifically sound judgements (Lamont 2009; Merton 1973). Consequently, scholars of peer review either have concentrated on scientists who independently make written evaluations or have focused on the exchange of arguments in small groups deliberating and judging the merit of research proposals. In Derrick's (2018: 8) words: "Deliberation is the hallmark of the peer-review process. Through deliberation, the panel develops a consensus by mitigating adjustments in the individual perspectives of panellists". In this regard, various studies (Huutoniemi 2012; Lamont 2009; Langfeldt 2001, 2004; Olbrecht and Bornmann 2010; Roumbanis 2017, 2022) closely examined the outcome of evaluation processes and how biases, group effects, and indeterminacy cause deviation in scientifically reasoned assessments. Taking fairness and impartiality as the baseline, other scholars persistently revealed empirically low agreement rates, unreasonable disadvantages, and conservative judgments (Ayoubi et al. 2021; Boudreau et al. 2016; Pier et al. 2018).

Such approaches support the view that peer review should be a well-reasoned exchange of scientifically grounded arguments between scientists. In this perspective, panel discussions sometimes go astray due to tacit side agreements, biases, and group effects. Within this framework, the ideal review process would only consist of scientific deliberations leading to clear and determinate funding recommendations. Any divergent (extraneous) influences in the form of emotions, group dynamics, and informal arrangements on the contrary must be avoided.

This idealistic view of peer review is omnipresent despite the contrary experiences of reviewers and the fact that observers of evaluation processes frequently report about limits of rationally made judgments and about assessments based on gut feelings. emotions, and strategic considerations (Derrick 2018; Huutoniemi 2012; Lamont 2009; Roumbanis 2017, 2022; Thorngate et al. 2009). Most scholars problematize such occurrences in peer review processes as leading to biased and unfair decisions. There are only few (Derrick 2018; Huutoniemi 2012; Lamont 2009; Lamont and Huutoniemi 2011; Raclaw and Ford 2015, 2017) who argue sporadically that informal exchanges, mitigating behavior, and emotion work contribute significantly to the success of research evaluations. New findings (Brunet and Müller 2024; Derrick and Bayley 2022; Paruschke et al. 2023; Peterson and Husu 2023) support this understanding of review procedures as a process that involves more than an exchange of scientific arguments to prepare the ground for decision-making. These studies not only report about bias, illegitimate behavior as well as supportive moves and practices but also call for a concept that takes into account such behavior as constitutive of panel meetings that evaluate research grants.

In this essay, I take these observed instances of strategic, informal, and nonverbal communication that occur before, during, and after rational deliberations about the

merit of applications to introduce a comprehensive concept for analyzing review panel processes. It is argued that evaluation processes, on the one hand, go from indeterminacy (vague, preliminary justifications) to determinacy (substantiated recommendations), while on the other, these processes are neither straightforward nor end in complete determinacy. Rather, each review panel goes through different phases (Derrick 2018; Lamont 2009; Paruschke et al. 2023), including formal as well as informal interaction orders, and while substantiated decisions are often documented. persistent vagueness is not. This vagueness has to do with certain occurrences, such as divergent views or tacit negotiations when determinate justifications are produced, but not with unclear grant proposals, the documented lack of details in the description of the methodology, or questionable feasibility. While this contextual vagueness of consensus formation is mainly associated with actions that push deliberations in certain directions, I propose that indeterminacy in the form of informal talk, emotion work, and mitigating behavior is also constitutive of review panel processes and has to be considered to fully grasp how review panel procedures evolve and produce funding recommendations. Only if one considers both determinacy and indeterminacy in review processes an effective research evaluation can be organized.

The essay starts with research about peer review processes and pays special attention to reports about phenomena that are not strictly associated with scientific reasoning and deliberation. Taking informal exchanges, mitigating behavior, and emotion work in review processes seriously, apart from illegitimate practices such as forming allies, manipulating other panelists, and taking shortcuts, I introduce the concept of the two constitutive intertwined layers of determinacy and indeterminacy. The paper discusses and closes with theoretical implications and practical consequences of this approach.

Empirical studies on peer review panels

Many investigations of peer review focus on written reviews, their wording, and scoring (Ayoubi et al. 2021; Kaatz et al. 2015; Witteman et al. 2019). Other studies closely examine evaluation processes involving authors, editors, and external reviewers to select scientific papers for publication (Hirschauer 2010; Kaltenbrunner et al. 2022; Myers 1989) and scientific committees that deliberate about and judge the merit of grant proposals (Abma-Schouten et al. 2023; Lamont 2009; Langfeldt 2001, 2004; Roumbanis 2017, 2022).

Research on the characteristics of evaluation panels and how they affect the decisionmaking process reveal that organizational constraints, such as the number of grant proposals, the chosen rating procedure, power imbalances, group settings, and the financial scope of a funding program, shape panelists' deliberations. However, there are two strands of research. The first (e.g., Derrick 2018; Huutoniemi 2012; Lamont 2009; Lamont and Huutoniemi 2011) is concerned with entire panel processes that assess the scientific quality of grant proposals. These scholars concentrate on how disciplinary repertoires, evaluative cultures, and situational definitions of assessment criteria influence consensus formations. They aim at the understanding and explanation of how collective judgments are reached despite epistemological and disciplinary differences. The second strand (e.g., Langfeldt 2001, 2004; Roumbanis 2017, 2022; Thorngate et al. 2009) investigates review panel meetings by comparing official guidelines, assessment criteria and proposed evaluation procedures with actually observable practices. It becomes evident that panelists in review processes vary in their use of criteria, that they practice task division, make concessions, or change, after tiring hours, from intense debates on proposals' merits to routines of selection and tacit compromises in the case of disagreements. Discussions of these findings predominantly circle around the peer review's intrinsic bias and its potential drawbacks in terms of financing unconventional research, small research institutes, and vulnerable groups of scientists, such as young and female researchers.

While the first strand of studies shows that under such conditions collective judgments can be achieved the second underpins—what Reinhart and Schendzielorz (2024) call—the 'Deficit Model' of peer review. Its dominance in scientific debates finally undermines the credibility of peer review in science and society more generally. In current debates about peer review, some scholars (Franzoni and Stephan 2023; Peterson and Husu 2023; Roumbanis 2019; Shaw 2023) therefore oppose such a mechanism of quality control because they relate its fairness and effectiveness to pure scientific deliberations about originality, feasibility, and potential impact. Subsequently, they discuss measures to either increase procedural purity or to abolish it all together.

Against this view, other scholars examine, among others, informal exchanges, mitigating behavior, and emotion work in panel meetings and mention their importance for successful and effective review processes (Derrick 2018; Huutoniemi 2012; Lamont 2009; Lamont and Huutoniemi 2011; Raclaw and Ford 2015, 2017). These studies frequently show if panelists discuss research proposals, they usually have incomplete information, have expertise in certain fields but not in others, and act under time pressures, which requires tactics and strategies to cope with conflicts and uncertainties. Under such conditions, a conversational analysis of panel interactions (Raclaw and Ford 2015), for example, shows how panel members use pauses, brief summaries (gist formulation), and implicatures (upshot formulation) to indirectly articulate a disagreement or a weak agreement. Other studies show that pre-meeting talks in general have a positive effect on the social inclusion of panel members (Yoerger et al. 2015) and laughter on dissolving divergent positions (Raclaw and Ford 2017). Such observations coincide with Lamont's (Lamont 2009; Lamont and Huutoniemi 2011) investigations of interdisciplinary panels and how the emotion work of panel members helps to manage diverging judgments and to settle conflicts (see also Mallard et al. 2009; Parker and Hackett 2014).

Recent research on review panels during the COVID-19 pandemic offers further insights into the importance of opportunities for panel members to exchange scientific arguments and for mitigating behavior, informal talks, and emotion work (Brunet and Müller 2024; Derrick and Bayley 2022; Paruschke et al. 2023; Peterson and Husu 2023). These studies empirically examined how panelists experienced the replacement of face-to-face meetings with video conference formats and its effect on evaluation processes. In line with previous research (Carpenter et al. 2015; Gallo et al. 2013; Pier et al. 2017) and looking back on lockdown restrictions, interviewed panelists

were mostly positive about remote gatherings via digital devices. From their point of view, they were able to continue with evaluations. However, they also described interactions in video conferences as more exhaustive and less dynamic and vivid compared to meetings in person. Especially in regard to evaluations of collaborative research centers and research training groups (Paruschke et al. 2023), they additionally complained about missing opportunities to collect impressions on-site, to have small talk with evaluated scientists, and to chat personally with other panel members about their research.

From an analytical perspective, one could say studies of peer review processes approach these procedures differently-from taking a certain aspect to the whole picture. There is research that concentrates on particular phenomena, such as the agreement between reviewers, their wording and scoring, and how it affects assessments and review outcomes (e.g., Ayoubi et al. 2021; Kaatz et al. 2015; Langfeldt 2001). These foci, of course, are important to disclose discrepancies and limits of peer review on grant proposals. They help to identify factors and areas that need further scrutinization and amendments. However, they also restrict investigation if informal exchanges, mitigating behavior, and emotion work are blamed for manipulating review outcomes. Another group of investigations is more inclusive by observing and studying the entire review panel procedure from its first gathering to the point when all members leave for their homes (Derrick 2018; Lamont 2009; Paruschke et al. 2023). In such studies in particular, one learns that in certain situations mitigating behavior, informal interaction, and emotion work are constitutive for effective review processes. There are various illustrative examples but no theoretical concept that brings rational deliberations and informal talks, emotion work, and mitigating behavior together as intertwined and constitutive layers of review panels.

The next section provides more detailed insights from previous studies about productive uses of emotion work, informal talk, and mitigating behavior. Based on these findings, I offer my theoretical framework of the intertwined determinacy and indeterminacy of peer review processes.

Looking at instances of indeterminacy in review panel meetings

There is an impressive list of drawbacks concerning peer review processes including gender bias, risk-avoiding grant allocation (conservatism), preferring research in one's own field of expertise (homophily), privileging large, esteemed research institutes (cronyism), favoring colleagues from the same research network (nepotism) and so on. Scholars explain these phenomena in review panel meetings with power imbalances as being analogous to a 'tug-of-war' fight (Derrick 2018), micro politics (Peterson and Husu 2023), organizational constraints (Langfeldt 2001), group dynamics (Derrick 2018; Olbrecht and Bornmann 2010) and collective anchoring effects (Roumbanis 2017). In contrast and apart from highly standardized procedures, clearly defined assessment criteria and ranking scopes, it is challenging to name constitutive factors in review panel processes which are more informal. Lamont's (2009) and Huuitiumi's (2012) works on interdisciplinary grant panel meetings, however, are valuable sources for instances on informal talks, emotion work, and

mitigating behavior. Both authors describe and characterize in detail disciplinary differences in deliberating and assessing the merit of applications and how criteria are used contextually. However, while these authors are primarily concerned with the intersubjective construction of consensus, this paper concentrates on illustrations in their work about situations of incommensurability and conflicts and how program officers and evaluators mitigate them and repair relationships among panelists.

Readers learn about the latter, for example, in Lamont's (2009) chapter on customary rules that are spelled out during panel meetings. One of these rules is that panelists respect the expertise of the other panel members. Interacting respectfully means they control their responses and emotions. In this context, Lamont (2009) observed a "consistently respectful tone towards one another" that "diffuses the potential for frictions and tensions that could hinder decision making, and helps create an amicable environment" (2009: 120). She also mentions contentious situations of explicit disagreement between evaluators and when panelists communicate their disparity in a more implicit, nonverbal form (e.g., "[f]rowning, rolling one's eyes, sighing, blushing", Lamont 2009: 140). As these occurrences can affect the group's ability to reach consensus, she describes in detail how panel members use emotion work and small talk to constitute a group that would respect each panelist's expertise or how "program officers, panel chairs, and some panelists engage in 'emotion work,' helping their colleagues save face even after defeat and reintegrating them into the group" (Lamont 2009: 140). On the following pages in her book, it becomes apparent that such forms of communication are also effective for integrating new panel members, handling discussions about incommensurable flaws in proposals, and keeping panel members motivated.

Respectful interaction was also observed by Huutoniemi (2012). In the observed interdisciplinary panels, she noticed an awareness of the expertise of other panelists and "academic politeness" when handling disputes on proposals. She also mentions the informal strategy to postpone discussions to a later point:

This became evident during a series of disagreements between two panel members whose opinions on several case study proposals were far apart. Both were experts in case-study methodology, but their theoretical backgrounds diverged. During a private discussion over breakfast, they came to an agreement concerning where their criteria of evaluation could overlap. (Huutoniemi 2012: 913)

According to Huutoniemi (2012), in this way panelists "tried to settle disagreements through mutual learning, compromising, or simply by trusting in each other's integrity and intuition" (912). Moreover, one can say returning back to these proposals towards the end of a meeting or after a pause left space for informal talk before the evaluators reexamined proposals with diverging assessments.

As a further example of group peer review processes that lighten the burden of evaluation, Derrick (2018) pointed to the effect of the established "committee culture" in the evaluation processes of the UK's Research Excellence Framework. She argues

that based on mutual respect for the other panelists after working together for several rounds, a strong committee culture produces a collectively shared definition of the proposal's quality that makes the process more efficient and helps to deal with "a large volume of submissions, a diverse panel membership and ambiguous objects" (Derrick 2018: 174). The author also emphasizes that opportunities for panelists to interchange between evaluation sessions help to increase group cohesiveness.

Support for these findings also provide Brunet and Müller (2024) who elaborate on emotions and how they enable reviewers to form their judgments. Based on interviews with European Research Council (ERC) panelists, they demonstrate that implicit norms define "how emotions should be expressed and experienced in different situations" (168) in evaluation processes. Both authors describe distinct feeling rules for different evaluative settings. With regard to in-person ERC panels, for example, they mention a feeling rule that "prohibits anger and supports respect for other reviewers' opinions" (169) and, as such, actively regulates the expression of emotions. The interviewed panelists state that confronted with persistent disagreements and conflicts over a proposal's qualities, panel chairs seize measures "to set up a cooperative, friendly, and joyful atmosphere between reviewers" that include activities such as "de-escalating conflicts as soon as they start[...], and fostering a 'team spirit' between panel members through organizing non-evaluation group activities (e.g., travel, housing, meals)" (180). Panelists in return comply with the feeling rule of respect and use their expression of emotions for different purposes. According to Brunet and Müller (2024),

they could apply three strategies of emotional management: over-emphasizing specific emotions, contextualizing the emotions expressed by others, and being simultaneously emotionally engaged and disengaged from the proposals (180).

The importance of informal talk and mitigating behavior, in addition, are issues in recent studies about peer review processes that moved from meetings in person to video conferences. Derrick and Bayley (2022), for example, warn that panelists in virtual settings have fewer opportunities to air disagreement and to settle conflicts because specific cues are likely to be reduced: "More so than ever the non-verbal, sometimes unconscious heuristics, are essential to enable (large) evaluation panels navigate and assess a complex and ambiguous object such as Impact" (95). Peterson and Husu (2023) also mention that participants of interdisciplinary review panels highlight the importance of informal talks that they are missing in virtual settings. In face-to-face meetings, they make use of informal talk to align with other panelists, particularly when proposals are outside of their core areas of expertise. An interviewed panelist, for example, said about talks during breaks:

It gives people a chance to formulate arguments more ... Because you talk [to other panelists] about whether this area is difficult or not. If you're not really into it [the specific area of the proposal], then you get a chance to understand what they think is difficult in that area. (3-P1-M in: Peterson and Husu 2023: 378)

When examining the impact of moving the evaluation of collaborative research centers and research training groups from meetings in person to video conferences, Paruschke et al. (2023) looked at the way this change affected different interaction orders (Goffman) in the peer review process. Distinguishing between formal, less formal, and situational interaction orders, they found that panelists in virtual settings often sorely missed opportunities for informal exchanges. From their views, formally organized deliberations worked well but were less vivid. They left enough space for reasoned exchange to produce a substantiated and well-reasoned funding recommendation. However, compared to meetings in person, it was more difficult to gain impressions of the context, to mitigate contentious situations, and to settle conflicts through talks in less formal interaction orders. In this regard, interviewees said:

That should not be underestimated, that is in any case also part of the evaluation, where then often information is obtained in private conversations, which one did not get in the large round, or where one asks again questions, of which one thinks, for the large round they are not so suitable, for it they go too much into detail or for it they are too problematic [...] and one puts then these over the canapés so to speak. (05 R, 26 in: Paruschke et al. (2023: 398), translated by the author)

And then also to get to know this working group [evaluated scientists] well, to see a little bit not only the specifications or ideas that are announced in written documents, but maybe also the attitude of the people and how they deal with these specifications and experience them. So, these conversations I think are essential. (18 G, 24 in: Paruschke et al. (2023: 398), translated by the author)

In a nutshell, different studies on peer review processes indicate that instances of indeterminacy such as informal talk, emotion work, and mitigating behavior are essential for evaluations to run smoothly and in such a way that all relevant aspects are included. Of course, allowing such extraneous influences opens the door to form allies, to manipulate other panelists, and to take shortcuts. However, if we only problematize these influences, there is the risk to lose track of their constitutional elements in review procedures. Moreover, it draws a picture of evaluation processes that are not perfect due to such extraneous influences. Even Lamont (2009), who disclosed how emotion work and informal talk can be effective, concludes that review panels are "an imperfect but satisfactory system" with "[s]trategic voting, horse-trading, self-interest, and idiosyncratic and inconsistent criteria all are unavoidable parts of *the equation*" (Lamont 2009: 156; highlighted by the author). In her "equation," formally organized deliberations are just one part to produce recommendations collectively. The other part are extraneous influences that one has to take into consideration to comprehend the outcome in this equation.

The equation analogy is not a key concept in Lamont's (2009) seminal and comprehensive work on consensus formation in review panels but with this model, one could explain why panelists are forced to deliberate pragmatically about the merit of proposals, following customary rules (e.g., deferring to expertise and respect for each other, bracketing self-interest, promoting methodological pluralism), and evaluating grant proposals relationally. In this equation-concept, nonetheless, extraneous influences are also reduced to be the factor that leads to imperfectly produced

recommendations. In a statistical equation, they would represent the potential error to make false decisions and that interpretations are correct only to a certain probability. This reminds us that fully valid decisions cannot be made due to either statistical errors or extraneous influences. It might explain panelists' behavior, but at the same time the equation-concept plays down that behavior and interactions outside of rational deliberations can have positive effects, and that they are an important factor to formulate funding recommendations. Informal talks, emotion work, and mitigating behavior are particularly pivotal in review processes to handle ambiguous objects, disagreements, and conflicts. To keep both aspects vivid, I differentiate between the contrarieties of determinacy and indeterminacy as constitutive layers of peer review processes. Determinacy represents all criteria, formal procedures, and reasoned deliberations. Indeterminacy stands for all inevitable informal acts and situations in peer review processes. Although contrarieties, I argue that collectively produced determinate decisions can only be comprehended by taking the complementary indeterminacy into account. In this regard, instead of Lamont's equation-concept, I draw on Boehm's (2007) picture theory and his concepts of how humans make sense of pictures.

The unity of contrarieties (in pictures)

According to the German picture theorist Gottfried Boehm (2007), to get to the meaning of a picture, one needs to reconstruct the iconic difference. Its meaning cannot be reconstructed sequentially like in text by going from word, phrase, and sentence to the next but by relating the visual elements to each other. In a picture, all visual elements are rather present simultaneously. Boehm says a meaningful image thus arises when something is brought into form in seeing. An observer produces meaning when focusing something optically in the indeterminate (against the horizon) and if it stands out by detaching it from the background. In other words, contrarieties are simultaneously present in pictures and constitutive for its meaning.

For Boehm (2007), paintings by Claude Monet are exemplary for pictural meanings that go beyond pure representation of determinable objects. If we look, for instance, at his painting of the Cathedral of Rouen from 1894, we easily spot and identify what we expected to find: the outlines of a building that could be found in reality in the streets of Rouen. However, there is more present and visible in the painting. A continuum of colored light distorts the recognizable silhouette of the cathedral and produces atmosphere at the same time. In this regard, the unique pictural meaning is constituted by the interplay of the building's determinacy and indeterminacy. What especially counts for Monet is constitutive for most pictures. The relation of form (determinate object) and ground (indeterminate surrounding) structures the way onlookers make sense of them. There are, of course, various encounters in art history to either represent objects in high contrasts and sharpness or dissolve any identifiable object in colors and shades. Both pictural forms have in common that viewers have difficulties to focus on something because determination is a process of marking a form or an area against an unmarked background.

In this respect, a further premise in pictural theory states that complete determinacy is unachievable. The indeterminacy of pictorial forms cannot be fully transferred into determinacy. There is always an "indeterminate remainder" (Wiesing 2018: 82) because understanding the depicted depends on its surrounding. In its contrariness, "the indeterminate and potential is the underlying ground of the specified" (Boehm 2007: 46, translated by the author). Accordingly, indeterminacy is not only the supporting foundation of the specified; rather, it has a different quality. It implies that fuzziness and vagueness in pictures are associated with potentiality and tension due to ambiguity. Vagueness in a picture, for example, adds atmosphere and mood, while specified figures provide clarity and guidance. As a consequence, picture theory takes indeterminacy seriously, assuming that it "creates initially the scope and potential that enables the specified to show *itself* and to show *something*" (Boehm 2007: 211, translated by the author). For the same reason, Wittgenstein (2003: 60, §71) argues that only blurred concepts (words) enable speakers to say something. Vagueness is needed to produce conceptual clarity and sharpness.

Borrowing the concept of simultaneously present contrarieties in pictures, I argue the intertwined contrarieties of determinacy and indeterminacy are also specific to peer review processes. It is characteristic of assessing scientific work collectively that it goes from indeterminacy to determinacy with both contrarieties present simultaneously at all phases. If one looks closer at whole panel review processes (as described in Derrick 2018; Huutoniemi 2012; Lamont 2009), panelists usually prepare for meetings by reading the assigned research proposals, collecting additional information about applicants, and making notes about a proposal's weaknesses and strengths. In joining the panel meeting in person, all participants bring their different perspectives and specialties and thus constitute an indeterminate situation. In the process of deliberating about the merit of research proposals, this indeterminacy ideally should turn into determinate judgments, which allow to make legitimate funding decisions.

However, the same literature reports that at the end of all these procedures, degrees of indeterminacy are still present. Full determinacy cannot be reached because in these processes panelists are "*calibrating* their individual senses of quality to a group standard in order to form a concerted evaluation" (Huutoniemi 2012: 910; Italics in original). In other studies, calibrating means "a pragmatic understanding of evaluation" (Lamont 2009: 125) or an "acceptable '*compromise*'" (Derrick 2018: 187; Italics in original). Derrick (2018: 200), for example, concludes in her book on the evaluation of Impact:

In peer review panels especially, individual panellists' standards must be calibrated and tensions among other panellists and their differing conceptualisations of Impact value are carefully managed in balancing acts of ongoing tugs-of-war requiring negotiation and compromise to reach a consensus.

From this perspective, indeterminacy cannot be turned into or replaced by full determinacy of produced judgments. Collectively made recommendations might

be plausible and reasonable but only because there often is no protocol on how panelists reached consensus, especially in the final stage of deliberations. This also becomes evident in cases when interviewed panel members reveal that their recommendations would be more reasonable if they had allowed "the panellists to go into more detail" (Huutoniemi 2012: 903) or that they feel uncomfortable with the panel's recommendation. About such feelings a professor in the life sciences and member of review panels, for example, said in a qualitative study about peer review and lottery (Philipps 2021) the following:

So as a rule, half of the proposals that come to the table are of an extremely good quality, where I would have great difficulty in saying that this is a reason to reject it. There is one thing somewhere in every proposal where you say okay, this could be improved, but it can be solved constructively in a process, so the proposal or the idea is not bad. [...] But you have the problem that you can't support the half that is good. And then there are the finer points, and that's where the big difficulties arise. What are the criteria? Is there already enough money? Is it the number of publications? Is it the impact of a paper? But how do I want to measure that? I may only be able to look at that after ten years. That's a very difficult issue and then you really go home from the panels and are actually always sad that you were only able to fund half of the good applications.

Apart from a panel's outcomes, there are profound clues that gaining determinacy in peer review processes is always accompanied by indeterminacy. This means that apart from discussing a proposal's originality, feasibility, and potential societal impact, panel members use informal talks, mitigating behavior, and emotion work in all phases of review processes (Lamont 2009; Paruschke et al. 2023; Roumbanis 2022; Thorngate et al. 2009). One reason is that panelists lack the time to consider all grant proposals to the same extent. They trust in their ability to exclude poor applications on the fly and concentrate their discussion on determining the merit of proposals close to the funding line (Brunet and Müller 2022). Often enough, indeterminacy reoccurs if panelists compare equally grant-worthy applications. Moreover, panelists' moods and impressions cannot be transferred into distinct scientific criteria. The social dynamics of such gatherings aim at clear funding recommendations, but they are not restricted to rational deliberations. The mood of the panelists and their impressions, prejudices, and contentions before, during, and between formally organized meetings affect the outcome (Derrick 2018; Lamont 2009; Paruschke et al. 2023; Thorngate et al. 2009).

Conclusion

Like art historians' caution concerning the constitutive potentiality of indeterminacy in pictures, researchers of panel review processes should intensify investigations of all forms of informal communication, practices, and emotion work related more or less to the panelists' situated and mutual scientific reasoning. Previously, scholars correctly observed indirect communication as door openers for strategic manipulations and biases, but—as I argue—they were badly advised to mark it as avoidable deviation.

On account of and in line with the foundations of picture theory, I propose that any determinate outcome of a panel review process cannot be understood unrelated to the indeterminate context. In the particular case of review processes, panelists' informal talks, emotion work, gist and upshot formulations, and so on during meetings shape and generate coherent judgments (Huisman 2001). Moreover, panelists sometimes even base their assessments on indeterminate impressions of the committee's atmosphere (Lamont 2009; Paruschke et al. 2023). Consequently, if indeterminacy is a constitutional part of panel review processes, we should learn more about it. A comprehensive understanding offers knowledge to better govern the evaluations of grant proposals. One might think particularly of conflicts that often cannot be solved through more rational arguments. A controlled move to indeterminate communication opens space for new perspectives, "to be emotional in appropriate ways" (Parker and Hackett 2014: 568), and allows a return to scientific reasoning of proposals' merit afterwards.

Furthermore, the understanding of peer review panel processes as constitutive intertwined layers of determinacy and indeterminacy goes beyond the idea of a progression from indeterminacy to determinacy. Thus, I argue that in panel meetings, indeterminacy in the form of informal talk, mitigating behavior, and emotion work constantly accompanies panelists' well-reasoned exchanges to produce determinate judgments. In fact, there are phases of indeterminate communication before and after phases of deliberations to assess proposals (Derrick 2018; Lamont 2009; Paruschke et al. 2023; Thorngate et al. 2009). Panelists informally talk on their way to and right before the meeting as well as in any pauses. They take their chance to chat about current research, to gossip, and to make small talk with less familiar or new panel members. All these behaviors and interactions are constitutive for collectively produced determinate recommendations.

Finally, the proposed concept's inclusive approach opens up a new field of research and integrates previous observations of review processes (Brunet and Müller 2024; Huutoniemi 2012; Lamont 2009; Langfeldt 2001, 2004; Roumbanis 2017, 2022; Thorngate et al. 2009) and foundations about the relevance of scientific criteria, values, and norms. It also relocates our understanding of social processes and local conditions of scientific peer review by shifting the focus away from actors' motivations, the definition of criteria, and strategic considerations in deliberations (Bornmann 2008; Gläser and Laudel 2005; Olbrecht and Bornmann 2010) to close investigations of interactions and practices as well as of various settings and situations in the entire review process starting from pre-meeting gatherings through to its closure. Examining indeterminate practices and contexts will offer new insights without questioning foundations such as that peer review assesses (a) intersubjective ideas and findings based on critical reference to previous and confirmed research, (b) that norms can lead scientists to conduct research according to the principles of "good" scientific practice, (c) that reviewers of scientific ideas and outcomes should be experts in respected fields, and (d) that selection in peer review only works under the condition of scarce resources (Bornmann 2008; Chubin and Hackett 1990; Lamont 2009; Langfeldt 2022). However, the theoretical framing of peer review processes as constitutive intertwined layers of determinacy and indeterminacy changes the perspectives on scientific evaluations. It is no longer about eliminating indeterminacy but about taking it seriously in order to better understand and manage scientific quality assessments.

Methodologically, the investigation of indeterminacy in review panel meetings will concentrate on contextual situations, informal practices, and forms of interactions that occur with scientific deliberations about the proposal's merits. In addition to interviews about panelists' understanding of assessment criteria, their enactment of feeling rules, and how they reach consensus in evaluation processes (cf. Brunet and Müller 2024; Huutoniemi 2012; Lamont 2009), the study of the aspects of indeterminacy will also examine closely recorded scripts of interaction and communication in panel meetings. Expanding on the work of Raclaw and Ford (2015), future research will index characteristic patterns of talk, interaction, and nonverbal communication in review panels.

A systematically reconstructed overview of aspects of indeterminacy in processes of producing determinate funding recommendations will provide information for all participants in review panel meetings. They will learn what can occur and how it might affect the evaluation process. This knowledge in particular will allow scholars, panelists, and panel chairs to distinguish between supportive informal practices in panel meetings and those that can endanger the review process. Equipped with such knowledge, they will be more sensitized to whether a deliberation needs extended negotiations, a restart, or shared reflection on the current illegitimate handling of the evaluation process.

A potential management implication may be a heightened sensibility of how funding recommendations are generated and whether it is legitimate. Informal reassurances or conversation among reviewers while eating canapés, for example, should not be a problem. They occur and are constitutive to formulate a collectively shared recommendation. At the same time, allowing such behavior is not a carte blanche to accept all recommendations that are well founded. Rather, care must be taken to ensure that determinacy is not bought by shortcuts, groupthink, and conservative decisions, especially if too many eligible applications are faced with a limited amount of funding. Overusing minor ambiguities and flaws in research proposals not only undermines the motivation of applicants but also increases mistrust in panel review procedures and thus their legitimacy in the long term. Questionably generated determinacy should therefore have consequences and prompt the chairpersons to a) obtain missing information, b) ask applicants for clarification, or c) place applications that are equally worthy of funding in a lottery bowl (Philipps 2022). Taking indeterminacy into account and handling it responsibly will help to increase the legitimacy of and trust in panel review processes.

Acknowledgements

I wish to thank the reviewers with their helpful feedback and Eva Barlösius for discussing and providing valuable thoughts on an earlier version of this paper.

Grant information

This research at the Leibniz University Hannover received funding from the German Bundesministerium für Bildung und Forschung (Grant 16PW18004).

References

- Abma-Schouten, Rebecca, et al. (2023), 'Evaluation of research proposals by peer review panels: broader panels for broader assessments?', *Science and Public Policy*, 50(4): 619–632.
- Ayoubi, Charles, Pezzoni, Michele, and Visentin, Fabiana (2021), 'Does it pay to do novel scicience? The selectivity patterns in science funding.', *Science and Public Policy*, 48(5): 635–648.
- Boehm, Gottfried (2007), Wie Bilder Sinn erzeugen. Berlin: Berlin University Press.
- Bornmann, Lutz (2008), 'Scientific Peer Review: An Analysis of the Peer Review Process from the Perspective of Sociology of Science Theories', *Human Architecture: Journal of the Sociology of Self-Knowledge*, 6(2): 23–38.
- Boudreau, Kevin J., et al. (2016), 'Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science', *Management Science*, 62(10): 2765-2783.
- Brunet, Lucas and Müller, Ruth (2022), 'Making the cut: How panel reviewers use evaluation devices to select applications at the European Research Council', *Research Evaluation*, 31(4): 486–497.
- --- (2024), 'The Feeling Rules of Peer Review: Defining, Displaying, and Managing Emotions in Evaluation for Research Funding', *Minerva*, 62(2): 167–192.
- Carpenter, Afton S., et al. (2015), 'A retrospective analysis of the effect of discussion in teleconference and face-to-face scientific peer-review panels', *BMJ Open*, 5(9): e009138.
- Chubin, Daryl E. and Hackett, Edward J. (1990), *Peerless science: Peer review and US science policy*. Albany: State University of New York Press.
- Derrick, Gemma (2018), *The Evaluators' Eye. Impact Assessment and Academic Peer Review*. Cham: Palgrave Macmillan.
- Derrick, Gemma and Bayley, Julie (2022), 'The Corona-Eye: Exploring the risks of COVID-19 on fair assessments of impact for REF2021', *Research Evaluation*, 31(1): 93–103.
- Franzoni, Chiara and Stephan, Paula (2023), 'Uncertainty and risk-taking in science: Meaning, measurement and management in peer review of research proposals', *Research Policy*, 52(3): 104706.
- Gallo, Stephen A., Carpenter, Afton S., and Glisson, Scott R. (2013), 'Teleconference versus Face-to-Face Scientific Peer Review of Grant Application: Effects on Review Outcomes', *PLOS ONE*, 8(8): e71693.
- Gläser, Jochen and Laudel, Grit (2005), 'Advantages and dangers of 'remote' peer evaluation', *Research Evaluation*, 14(3): 186–198.
- Hirschauer, Stefan (2010), 'Editorial Judgments: A Praxeology of 'Voting' in Peer Review', *Social Studies of Science*, 40(1): 71–103.

- Huisman, Marjan (2001), 'Decision-Making in Meetings as Talk-in-Interaction', *International Studies of Management & Organization*, 31(3): 69–90.
- Huutoniemi, Katri (2012), 'Communicating and compromising on disciplinary expertise in the peer review of research proposals', *Social Studies of Science*, 42(6): 897– 921.
- Kaatz, Anna, et al. (2015), 'A quantitative linguistic analysis of national institutes of health r01 application critiques from investigators at one institution', *Academic Medicine*, 90(1): 69–75.
- Kaltenbrunner, Wolfgang, et al. (2022), 'Changing publication practices and the typification of the journal article in science and technology studies', *Social Studies of Science*, 52(5): 758–782.
- Lamont, Michèle (2009), *How professors think: Inside the curious world of academic judgment*. Cambridge, Mass.; London: Harvard University Press.
- Lamont, Michèle and Huutoniemi, Katri (2011), 'Comparing Customary Rules of Fairness: Evaluative Practices in Various Types of Peer Review Panels', In: Charles Camic, Neil Gross, and Michèle Lamont (eds.), *Social Knowledge in the Making* (209–232. Chicago: The University of Chicago Press).
- Langfeldt, Liv (2001), 'The decision-making constraints and processes of grant peer review, and their effects on the review outcome', *Social Studies of Science*, 31(6): 820–841.
- --- (2004), 'Expert panels evaluating research: decision-making and sources of bias', *Research Evaluation,* 13(1): 51–62.
- --- (2022), 'The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome', In: Eva Forsberg, et al. (eds.), *Peer review in an Era of Evaluation: Understanding the Practice of Gatekeeping in Academia* (297–326. Cham: Springer).
- Mallard, Grégoire, Lamont, Michèle, and Guetzkow, Joshua (2009), 'Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review', *Science, Technology, & Human Values,* 34(5): 573–606.
- Merton, Robert K. (1973), *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.
- Myers, Greg (1989), 'The pragmatics of politeness in scientific articles', *Applied Linguistics*, 10(1): 1–35.
- Olbrecht, Meike and Bornmann, Lutz (2010), 'Panel peer review of grant applications: what do we know from research in social psychology on judgment and decisionmaking in groups?', *Research Evaluation*, 19(4): 293–304.
- Parker, John N. and Hackett, Edward J. (2014), 'The Sociology of Science and Emotions', In: Jan E. Stets and Jonathan H. Turner (eds.), *Handbook of the Sociology of Emotions: Volume II* (549–572. Dordrecht: Springer Netherlands).
- Paruschke, Laura, Barlösius, Eva, and Philipps, Axel (2023), '(Fehlende) Anwesenheit bei Begutachtungsprozessen: Zum Wandel der sozialen Interaktionen im virtuellen Format', *Zeitschrift für Soziologie*, 52(4): 391–403.
- Peterson, Helen and Husu, Liisa (2023), 'Online panel work through a gender lens: implications of digital peer review meetings', *Science and Public Policy*, 50/3: 371–381.

- Philipps, Axel (2021), 'Science rules! A qualitative study of scientists' approaches to grant lottery', *Research Evaluation*, 30(1): 102–111.
- --- (2022), 'Research funding randomly allocated? A survey of scientists' views on peer review and lottery', *Science and Public Policy*, 49(3): 365–377.
- Pier, Elizabeth L., et al. (2017), 'Videoconferencing in Peer Review: Exploring Differences in Efficiency and Outcomes', In: B. K. Smith, et al. (eds.), *Making a Difference: Prioritizing Equity and Access in CSCL* (2. Philadelphia, PA: International Society of the Learning Sciences).
- Pier, Elizabeth L., et al. (2018), 'Low agreement among reviewers evaluating the same NIH grant applications', *Proceedings of the National Academy of Sciences of the United States of America*, 115(12): 2952–2957.
- Raclaw, Joshua and Ford, Cecilia E. (2015), 'Meetings as Interactional Achievements', In: Joseph A. Allen, Nale Lehmann-Willenbrock, and Steven G. Rogelberg (eds.), *The Cambridge Handbook of Meeting Science* (247–276. Cambridge: Cambridge University Press).
- --- (2017), 'Laughter and the management of divergent positions in peer review interactions', *Journal of Pragmatics*, 113: 1–15.
- Reinhart, Martin and Schendzielorz, Cornelia (2024), 'Peer-review procedures as practice, decision, and governance—the road to theories of peer review', *Science and Public Policy*.
- Roumbanis, Lambros (2017), 'Academic judgments under uncertainty: A study of collective anchoring effects in Swedish Research Council panel groups', *Social Studies of Science*, 47(1): 95–116.
- --- (2019), 'Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-making Mechanisms for Allocation of Research Grants', *Science Technology and Human Values*, 44(6): 994–1019.
- --- (2022), 'Disagreement and Agonistic Chance in Peer Review', *Science Technology and Human Values*, 47(6): 1302–33.
- Shaw, Jamie (2023), 'Peer Review, Innovation, and Predicting the Future of Science: The Scope of Lotteries in Science Funding Policy', *Philosophy of Science*, 90: 1297–1306.
- Thorngate, Warren, Dawes, Robyn M., and Foddy, Margaret (2009), *Judging Merit*. New York: Psychology Press.
- Wiesing, Lambert (2018), *The visibility of the image. History and perspectives of formal aesthetics*. London: Bloomsbury Academic.
- Witteman, Holly O., et al. (2019), 'Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency', *The Lancet*, 393(10171): 531–540.
- Wittgenstein, Ludwig (2003), *Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp.
- Yoerger, Michael A., Francis, Kyle, and Allen, Joseph A. (2015), 'So Much More than "Chit-Chat": A Closer Look at Premeeting Talk', In: Joseph A. Allen, Nale Lehmann-Willenbrock, and Steven G. Rogelberg (eds.), *The Cambridge Handbook of Meeting Science* (Cambridge Handbooks in Psychology, 153–174. Cambridge: Cambridge University Press).