

Selling science: optimizing the research funding evaluation and decision process

Selling science

1

Claartje J. Vinkenburg

Independent Expert, Amsterdam, The Netherlands

Carolin Ossenkop

Connectify, Wijchen, The Netherlands and

Institute for Management Research, Radboud University Nijmegen, Nijmegen, The Netherlands, and

Helene Schiffbaenker

Joanneum Research Forschungsgesellschaft mbH, Graz, Austria

Received 28 January 2021
Revised 14 July 2021
Accepted 23 July 2021

Abstract

Purpose – In this contribution to EDI's professional insights, the authors develop practical and evidence-based recommendations that are developed for bias mitigation, discretion elimination and process optimization in panel evaluations and decisions in research funding. An analysis is made of how the expectation of "selling science" adds layers of complexity to the evaluation and decision process. The insights are relevant for optimization of similar processes, including publication, recruitment and selection, tenure and promotion.

Design/methodology/approach – The recommendations are informed by experiences and evidence from commissioned projects with European research funding organizations. The authors distinguish between three aspects of the evaluation process: written applications, enacted performance and group dynamics. Vignettes are provided to set the stage for the analysis of how bias and (lack of) fit to an ideal image makes it easier for some than for others to be funded.

Findings – In research funding decisions, (over)selling science is expected but creates shifting standards for evaluation, resulting in a narrow band of acceptable behavior for applicants. In the authors' recommendations, research funding organizations, evaluators and panel chairs will find practical ideas and levers for process optimization, standardization and customization, in terms of awareness, accountability, biased language, criteria, structure and time.

Originality/value – Showing how "selling science" in research funding adds to the cumulative disadvantage of bias, the authors offer design specifications for interventions to mitigate the negative effects of bias on evaluations and decisions, improve selection habits, eliminate discretion and create a more inclusive process.

Keywords Research funding, Panel evaluation, Decision making, Bias mitigation, Discretion elimination, Process optimization, Inclusion, Selling science

Paper type Viewpoint

Selling science in research funding

In this contribution to *EDI's Professional Insights*, we develop recommendations for improving evaluation and decision processes involving groups such as panels and committees in research funding. Our insights are relevant for process optimization in similar evaluative decision processes, including publication, recruitment and selection, tenure and promotion in higher education and research settings. Our recommendations are informed by evidence and experiences from commissioned projects in the context of European



© Claartje J. Vinkenburg, Carolin Ossenkop and Helene Schiffbaenker. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Equality, Diversity and Inclusion:
An International Journal
Vol. 41 No. 9, 2022
pp. 1-14
Emerald Publishing Limited
2040-7149
DOI:10.1108/EDI-01-2021-0028

research funding decisions, through interviewing panelists and observing panels. These projects were initiated with the purpose of uncovering and mitigating bias (i.e. cognitive distortion) in the evaluation process.

The evaluation and decision process involved in research funding, similar to other selection decisions for prestigious and scarce positions, can be characterized as imperfect or suboptimal (Vinkenburg *et al.*, 2014; van den Besselaar *et al.*, 2018). Room for bias occurs when evaluators are given discretionary space in evaluating performance and potential. In these kinds of complex decision processes, with limited time and inherent intransitivity, the volatile application of ambiguous criteria such as excellence, groundbreaking or “high risk high gain” induces the occurrence of bias. The stronger evaluators believe that the research system is meritocratic, the larger the room for bias to affect the outcome of the evaluation (Régner *et al.*, 2019). While bias is often implicit, it becomes audible and readable through language and thus apparent to trained observers.

Research funding is both an indicator of and a (increasingly) necessary requirement for career success in academia (Bloch *et al.*, 2014; Bol *et al.*, 2018). The expectation of “selling science” as a critical component of research funding decisions became apparent in interviews we held on and observations we made of evaluation processes in the past decade. Competition implies that evaluators in their role as panelists or reviewers are asked to compare, weigh and decide whether to buy the ideas that an individual applicant, research team or consortium tries to sell them. While the ability to “sell science” is not a formal selection criterion and/or stage of the process, it emerges as a tacit criterion during the process and is a distinctive part of the narrative that evaluators use to describe why some are more or less likely to be selected and ultimately funded than others. We explore this phenomenon to provide professional insights and recommendations for process optimization and discretion elimination in evaluation.

Evaluators bring their own, often implicit, image of the “ideal academic” (or ideal scientist, career, team or consortium) into the process (Leslie *et al.*, 2015). This ideal image or prototype is influenced by the discipline of the evaluator and the nature of the grant scheme, but the degree to which the applicant (or team) fits the stereotypical image determines their chances of success (Bickmore *et al.*, 2018). Selling science with its connotations of self-promotion, assertiveness and confidence produces lack of fit for those who are held to injunctive norms of modesty or otherwise do not fit the part (Fox *et al.*, 2017). The expectation of embodying and enacting the ideal image through situated performance in presentations and interviews makes fit or lack thereof almost tangible.

Prescriptive gender stereotypes or injunctive norms produce a narrow band of acceptable behavior for women when applying for prestigious roles or positions. When women act according to stereotypical expectations, they remain invisible; when women actively sell their science, they experience backlash. Therefore, the research funding evaluation and decision process is a particular and stark example of women’s perceived lack of fit with the stereotypical or ideal academic (Bleijenbergh *et al.*, 2013; Heilman *et al.*, 2015; van Veelen and Derks, 2020). As explained by Biernat *et al.* (2020), shifting standards imply that when men stereotypically are expected to have more of an attribute than women, individual women are judged on that attribute relative to (lower) expectations for women, and men relative to (higher) expectations for men. Gender frames both *viewing* and *valuing* difference when making evaluations about who is truly exceptional, creating a double advantage for men (Correll *et al.*, 2020). Acknowledging intersections of bias and stereotypes based on gender with those on other irrelevant but visible characteristics such as ethnicity or race, class, age and physical ability further illuminates how research funding decisions fuel cumulative advantage for White men and the inherent homogenization of the research system.

To illustrate the effects of bias on the evaluation process, we distinguish three different *aspects* of the process in research funding decisions, namely evaluating written information in

applications, evaluating enacted performance in interviews and presentations, and group dynamics within the panel. We introduce each aspect of the evaluation process by sharing an evidence-based *vignette* informed by our observations and experiences in supporting research funding organizations as external experts. Understanding the vignettes as illustrations of how bias affects the evaluation process from the perspective of different stakeholders, we set the stage for our analysis and for developing actionable recommendations on how to mitigate bias and reduce discretionary space in the evaluation process. We analyze how bias occurs in each of these aspects and explain how selling science adds a layer of complexity to an already cognitively demanding situation. In the analysis, we use examples from *interviews* with panelists [1] held in the context of a commissioned project on bias in research funding (van den Besselaar *et al.*, 2018), and we refer to research evidence on similar evaluations. We then provide *recommendations* to optimize the process for each aspect in particular and for the overall process. Raising awareness is important but not enough to mitigate the effects of bias and change evaluation and decision habits. Going beyond awareness, we elaborate on accountability (A), biased language (B), criteria (C), structure (S) and time (T). Our recommendations are listed in [Tables 1–4](#) below, following the same AABCST logic. Through process optimization and discretion elimination, we contribute to diversity and inclusion in higher education and research settings.

The evaluation process in research funding: the perils of selling science

Bias makes its way into the evaluation process as it affects panelists' implicit associations, explicit expectations and behaviors. Implicit notions of the "ideal academic" and the "ideal career" affect how panelists read, view, hear, question, describe and discuss applications. In order to obtain research funding, it is necessary to sell science. This means that on paper, applicants have to write well and convincingly to be perceived as excellent. In person, applicants are expected to act confident to be perceived as competent. For panelists to recognize merit, they have to buy what is being sold. In principle, writing and acting in a confident manner have low predictive validity for future research performance. At the same time, evaluators often confuse confidence with competence. More generally speaking, criteria such as institutional prestige are at best proxies for research performance. Especially for early career researchers, being judged on (research) potential rather than past performance results in the benefit of the doubt going to those who fit the "ideal image" better than those who do not. Nonperformance factors and irrelevant personal characteristics contribute to perceived (lack of) fit, and thus, the shared expectation of selling science makes bias explicit.

Interestingly, some panelists refer to the notion of overselling, implying that some applicants by virtue of being highly skilled and even groomed in writing and presentation style raise expectations of future performance that they may or may not be able to meet. Some panels (or members) are more sensitized towards overselling than others, perhaps related to the degree to which certain disciplines have a stronger tradition of hiring external consultants and coaches in preparing applications. The question is then whether overselling gets "caught" or "bought", and whether as a strategy it benefits or harms some more than others.

Selling and overselling science may play out differently depending on the stage of the evaluation process. Some panelists view overselling as the norm and argue that underselling will result in a written application being rejected immediately. Other panelists argue that the interview and presentation provide the chance to really learn about an applicant's merits. The norm of overselling is challenged when applicants with a sound, but not oversold written application, are able to show their qualities in the presentation or interview. This pattern was reported for women applicants more than men. In contrast, some applicants who oversell in writing are not able to meet these expectations in enacted performance.

This pattern was reported for men applicants more than women. So *not* overselling in a written application can be an advantage for women especially, highlighting the narrow band of acceptable behavior.

Bias in the evaluation of written materials

4

Vignette 1 – evaluating written materials

“And I think: I will not let this one be yet another cycle of grant allocations!”

It was just the start of yet another cycle of grant allocation, I thought...

Just as with any other panel for grant allocation, I thought, based on CVs and research proposals, I will now receive the chair’s initial ranking of applicants to be invited for the interviews, I thought, about to open the email...

Just as with any other panel for grant allocation, as expected of me (*in my role as scientific officer of the panel*), I thought, I will send this list to all panelists, I thought, as preparation for the subsequent panel meeting set to officially determine the ranking of applicants, I thought...

Opening the email and quickly going through the list of names,
... wait a minute ...

"All men are ranked at the top, while all women are not?"

WAIT A MINUTE, I think...

When asked to evaluate written materials, typically a CV or track record and an idea or research proposal, the panelists’ implicit associations of the ideal application are made salient by the style, vocabulary and content of what is written (Markowitz, 2019). Panelists describe the writing as “assertive” or even an “aggressive” way of self-promotion, in terms of selling science and scientist as worth funding. This style is said to serve applicants (and especially men) well in terms of leading to high impact publications. Panelists assume that gender differences in writing style exist; they read more confidence, ambition and “fanciness” in men’s than in women’s written applications.

Interestingly, some panelists describe the written materials provided by women as better prepared and more realistic, sound and substantial than those prepared by men. They argue that falling for the (over-)confident and exaggerated writing style of men may serve as an exclusionary mechanism for women whose writing is perceived as humbler and more realistic. Overselling on paper is sometimes caught when an exceedingly well-written proposal using superlatives or buzzwords raises doubts among panelists in terms of who actually wrote the proposal. Poor writing obviously affects applicants’ chances but may also prompt panelists to look at it again and read through the lines to look for great ideas.

Implicit associations also affect how reviewers and panelists write about applicants and applications. Biased language in evaluation reports is found in terms of certain adjectives (e.g. superlatives), nouns (e.g. “leader”) and negations (e.g. “she does not have a bad CV”) and can be identified using linguistic software (Kaatz *et al.*, 2015). Some evaluation procedures require applicants to submit recommendation letters, which is another place where such linguistic bias (e.g. doubt raising) becomes apparent (Madera *et al.*, 2018) (see Table 1 for recommendations on evaluating written materials).

Evaluating written materials	
Awareness	Write your own comments
Accountability	Fill out forms completely
Biased language	Be aware of your ideal image
Criteria	Know the red flags that trigger your bias in a written text
Structure	Recognize gendered adjectives, nouns and pronouns
Time	Scan written text (comments and recommendations) using app or software
	Do this before opening the application files
	Explain scores using both inclusion and exclusion arguments
	Decide on and apply compensation rules (within and between criteria)
	Evaluate per criterion, not per application
	On each criterion
	On each application
	Check down and across the scoring form

Table 1.
Recommendations
evaluating written
materials

Vignette 2 – evaluating enactment (abridged) - unabridged version in supplementary materials

“#honestly...”

[...] I enter the room through the door held by the person,
who opened it a few moments ago.

I smile. I look around.

I try to get a glimpse of all the faces in the room in a split second.

I smile.

Do I smile too much?

They don't smile. Only one of them.

What does that mean? What do they think?

Will they like me? Will they like my idea?

Will they like what I say? Will they like how I say it?

Looking at their faces, I am afraid they...

I should stop thinking like that! STOP!

I felt...

STOP. FOCUS.

I arrive at the lectern. [...]

OK - SHOWTIME!

I take a deep breath.

I hear myself say: "Good morning everyone.

Thank you for giving me the opportunity
to present my proposal today."

They still look so serious.

Was that not the right opening words? What else
should I have said? Maybe it would have been better
to stick to what Rogier told me? STOP.

Now I already said it, I will just go on.

Maybe they don't really hear that begin-part
anyway. This was so quick anyway...

They will not pay attention to it

– or will they? STOP. FOCUS.

[...]

I feel myself behind the lectern; their eyes on me,
on my clothes, on my body, (on my pen?),
(my pen?)? on my body? On my clothes? On me?
behind the lectern...

... behind the lectern??? I have been standing

behind the lectern. I feel comfortable here, behind the lectern. [...]

– but should I be standing behind the lectern -
throughout my whole talk? What will they think?

Shouldn't I take the whole space that is there – in front of the slides?

Take the space? Own the space?

Like I have seen so many others do previously?

What did they say again in the prep-course for this occasion?

Take the space?

Own the space?

STOP. FOCUS.

I go on.[...]

Selling science is an enacted performance that requires a stage and an audience. The enactment in presentations and interviews is an opportunity for applicants to make panelists “buy” their application. The applicants appear and perform on a stage while they are watched by a group of evaluators who respond both nonverbally and verbally to what is enacted. In such a demanding atmosphere, this performance may appear as standing in front of a tribunal, which is not equally appreciated by and fitting for all applicants.

Evaluation of interview and presentation (enactment)	
Awareness	<p>Be aware of enactment</p> <p>Be aware that panelists actively affect enactment</p> <p>Hold each panelist accountable</p> <p>Be aware that confidence is not equal to competence</p> <p>Be aware of shifting standards</p> <p>Apply the same criteria to each applicant</p> <p>Ask the same questions to each applicant</p> <p>Spend equal amount of time</p>
Accountability	<p>Acknowledge that interviews/presentations stage require enactment (as in performing on stage or putting on a show)</p> <p>Acknowledge that the expectation to 'enact' the ideal image fits some applicants better than others</p> <p>Acknowledge that the quality of the performance has little predictive validity for the quality of the research</p> <p>Acknowledge that panelist are active audience members whose responses (smiling, affirming, discouraging, disapproving) have an effect on the enactment</p> <p>Write your own comments and fill out forms completely during the presentation and/or the interview, independently from the other panelists</p> <p>Be sensitive to biased language used by applicants (superlatives, pronouns such as I versus we, etc) as a sign of (over) coaching and (over) selling or of making oneself bigger or smaller than one is</p> <p>Be sensitive to shifting standards expressed in language</p> <p>Reduce your own use of biased language in your response to the enactment</p> <p>Be sensitive that the enactment itself is not being evaluated and scores on the formal criteria are colored by the enactment performance</p>
Biased language	<p>Use a standard list and order of questions for each applicant</p> <p>Use a scoring form/rubric to assess each applicant in the same way</p> <p>On each criterion</p> <p>On each application</p> <p>Check down and across the scoring form</p>
Criteria	
Structure	
Time	

Table 2.
Recommendations
evaluation of interview
and presentation
(enactment)

What sells well depends again on the norms in the discipline, but it also depends heavily on panelists' individual preferences and expectations. Panelists may not appreciate an applicant who dresses in a certain way, who has a voice perceived as too loud or not loud enough, who smiles too much or too little. As each panelist has an individual picture of the ideal applicant, each applicant's individual enactment may fit this image more or less.

Language is specifically relevant for identifying bias in enacted performance, in terms of language used by applicants as well as language used by panelists. Through the applicant's word use and language skills, panelists might be more or less impressed by and fond of the performance. An applicant who does not use the expected buzzwords for hot topics in one's discipline may receive lower support and less beneficial evaluations compared to the evaluation of an applicant whose vocabulary fits the ideal. The same goes for using self-promotional language in terms of superlatives, boasting and pronouns ("I" versus "we"). Applicants have limited opportunities to learn about and prepare for these (often informal or implicit) rules and preferences; transparent criteria might not or only vaguely exist.

When overselling is part of their enactment, applicants appear confident, and panelists can perceive and easily assess this confidence as competence. Indeed, panelists may have gendered expectations in terms of overselling: While acting (very) confident and assertive is clearly expected and appreciated in the enacted performance of men, this is less the case for women. Women are expected to act more modest and less confident and (for those sensitized to the issue) not to oversell. They may be assessed more positively when their enacted performance meets gendered norms for women, when panelists perceive them as well-prepared and nice. Enacted performance is particularly prone to the impact of shifting standards. For example, while being or appearing to be young is seen a sign of potential in men, it is seen as a sign of inexperience in women (Malmström *et al.*, 2017). These shifting standards become audible through the language used by panelists when posing questions to, giving feedback on or raising doubts about applicants during and after the performance (see Table 2 for recommendations on the evaluation of interview and presentation).

Bias due to group dynamics

Vignette 3 – group dynamics

“Taking a look behind the scenes”

The chair leans forward: “This stays confidential, right?”

I am quick to say: “Yes, yes. Of course.”, and the others on his panel nod, as expected.

He continues: “You know, if I have a favorite (*applicant*) – and usually, I know at least some of the guys on the panel quite well – I know exactly, who will think along the same lines and thus, who I need to give the word, first, (*to find support and set the tone in support of my own preference*)...”.

Essentially, each panel is a group of individuals. Groups and their members are sensitive to specific dynamics, both immediately present and/or gradually developing throughout the group process. Especially decision-making situations are prone to group dynamics, affecting

Group dynamics	
Awareness	<p>Make this explicit (know the benefits and flaws of group decision making)</p> <p>Make this explicit (structure reduces room for bias)</p> <p>Structure of "who speaks first" (e.g. rotation or novice rule); how to make final decisions</p> <p>Reconsider documents, be complete to reduce need for tacit and informal knowledge/assumptions</p> <p>Know that such complex, comparative evaluations that require a ranking suffer from inherent intransitivity: Paper, rock and scissors (i.e. no evident winners)</p> <p>Discuss and agree (e.g. filling in all relevant templates/forms)</p> <p>Use less biased language and call on each other when doing so</p> <p>Discuss each criterion prior to starting the evaluation round (prevents relying on unclear or flawed understanding)</p> <p>Discuss/make this explicit (e.g. weighing and compensation of criteria)</p> <p>Evaluate per criterion, not per candidate</p> <p>Provide both inclusion and exclusion arguments</p> <p>Discuss and agree upon evaluation structure</p> <p>Implement required changes (e.g. flip evaluation matrix to "per criterion")</p> <p>Applying evaluation structure means to also let discussion evolve accordingly (e.g. discuss per criterion, not per candidate)</p> <p>Introduce and agree on discussion structure (e.g. who speaks first/rotate speaking time per panelist, etc.)</p> <p>Seating arrangement, e.g. mix men and women, low and high status panelists around the table or in the meeting application</p> <p>Needs to be accepted by panelists, supporting role in meetings and coach to chair and scientific officer</p> <p>Prescribe the exact questions, order and panelist to ask the question across all panelists</p> <p>Plan this ahead of time and communicate and apply</p>
Accountability	
Biased language	
Application of criteria	<p>Understand panel as a group</p> <p>Introduce need for structure</p> <p>Discuss and agree on structure</p> <p>Make all rules and procedures explicit</p> <p>Be aware of intransitivity</p> <p>Everybody is accountable</p> <p>Sensitivity to language use</p> <p>Explicitly agree (or disagree) on relevant criteria</p> <p>Agree on how to apply criteria</p> <p>Agree and apply evaluation structure</p>
Structure (for panel)	<p>Align discussion with evaluation structure</p> <p>Apply discussion structure</p> <p>Optimally structure physical or virtual meeting environment</p> <p>Invite external participant expert observer</p> <p>Apply strict interview guide</p> <p>Spend equal time on all criteria and applications</p> <p>Reserve time for inclusion arguments</p> <p>Reserve time for calibration</p>
Structure (for interviews)	
Time	

Table 3.
Recommendations
group dynamics

the decision outcome substantially (King *et al.*, 2018). In research funding evaluation and decision-making, intersecting factors such as status hierarchy, power distribution and general conventions of interaction pave the way for decisions made in favor of some applicants in comparison to others. For instance, the convention of consistently giving the word to the most senior or highest status panelist first turns this panelist's impression and evaluation into an anchor for the rest of the discussion. Adopting a power lens illuminates how it is particularly challenging for less senior and lower status panelists to voice diverging impressions and evaluations in response. If worse comes to worst, such group dynamics may (unintentionally) lead to decision-making (re)producing systematic inequalities between members associated with particular social groups.

Particularly, the panel chair and scientific officer supporting the chair are game changers with respect to group dynamics. Thanks to their formal role, they can influence the group dynamics more structurally – for better or worse. Continuing on the previous example, usually, the chair asks panelists for their input to the discussion. As such, the chair can either embrace the general convention of asking the highest status panelist first, or the chair can organize the discussion differently (e.g. taking turns with whom speaks first). Thus, the chair can actively regulate the room for bias to affect decision-making.

The comparative and intransitive nature of panel decision-making, in which multiple applications are ranked on the basis of multiple criteria, leaves room for bias but also provides opportunities for process optimization in terms of explicitly operationalizing and weighing criteria and of deliberate calibration between and across ranked applicants (Vinkenburg, 2017). Again the possible impact of the chair on this part of process is quite large (see Table 3 for recommendations on group dynamics).

General recommendations

With these more detailed analyses of how bias affects three different but interrelated aspects of the evaluation and decision process, and with concrete recommendations on each aspect for panelists and chairs, we hope to have raised awareness (A) on how process optimization and discretion elimination can help eliminate the cumulative (dis)advantage of the requirement of selling science in research funding. Even if the room for improvement is marginal, changing selection habits and mitigating the effects of bias contribute to a more inclusive process and outcomes. It is important that all involved (panel members, scientific officers who support panels, and especially panel chairs) are held accountable (A) for their contribution to decision outcomes and are engaged in process optimization. Some research funding organizations allow for an appeal if the applicant (or a panelist) believes that bias has occurred and/or if insufficient or incorrect information is provided to justify a decision. These appeals are then considered in a double blind setting.

To provide insight into and offer practical recommendations for improving inclusive, non-biased language (B), a linguistic bias scan or app can be used on materials (e.g. guide for applicants), on written applications, reference letters and evaluation reports but also to support observations in panels. Emphasizing that bias becomes explicit and audible in language and giving evidence based, context specific examples of biased language provide a means to move away from general concerns panelists may have around political correctness. As language is evolving and fluid, showing the negative impact of language on evaluation is important, especially for internationally and/or disciplinary diverse panels.

To make the application of criteria (C) less volatile, it is important to consider and discuss the meaning and operationalization of vague criteria such as excellence or research potential and of rating scales (e.g. “beyond expectations”) among evaluators before moving on to the actual evaluations. Funders (or employers) are challenged to specify criteria in more detail, by

General recommendations	
Awareness	<p>Assign referee role to officer or observer</p> <p>Do not allow discussions about criteria, applications and outside formal meeting</p> <p>Sanction unwanted behavior/norm violations</p> <p>Expose the “false positive” consequences of overselling</p> <p>Mention it in call materials</p> <p>Provide training to evaluators (using video)</p> <p>Call bias when it is seen/heard/observed</p> <p>Expose the narrow band, broaden the ideal image and reduce lack of fit</p> <p>Scan written text (including call texts and evaluation reports) using software</p> <p>Use a linguistic bias checklist or app for observers</p> <p>Carefully consider using narrative CV formats (room for linguistic bias)</p> <p>Specify/disambiguate criteria more</p> <p>Discuss the meaning of criteria when training evaluators/panels first meet</p> <p>Demand using both inclusion and exclusion arguments</p> <p>Communicate compensation rules (within and between criteria)</p> <p>Build in calibration time and instructions</p> <p>Build in appeal on process and criteria application</p> <p>Wild idea: develop a criterion on the enacted performance as such</p> <p>Wild idea: build in a threshold for lottery and/or person-job fit</p> <p>Change forms and process to evaluate per criterion, not per applicant</p> <p>Develop and introduce instructions on discussion structure, interview guide, etc</p> <p>Split the evaluation of the CV/idea (in the process, the package or the panel)</p> <p>Drop or blind the evaluation of the CV</p> <p>Check and control time spend on each</p> <p>Lottery, thresholds plus lottery, veto and wild cards</p> <p>Virtual versus physical settings: differences, similarities, needs and challenges</p>
Accountability	<p>Raise bias awareness</p> <p>Change selection habits</p>
Biased language	<p>Address use of biased language</p>
Criteria	<p>Specify meaning of criteria</p>
Structure	<p>Apply criteria in equal measure to all applications</p>
Time	<p>Spend equal amounts of time</p>
Wild	<p>Wild ideas</p>

Table 4.
General recommendations and wild ideas

providing examples of how to operationalize criteria given a particular grant scheme or call for proposal. Acknowledging that evaluators are experts when it comes to the discipline or the science involved, it is important to stress that reducing ambiguity and promoting transparency contributes to the predictive validity of the evaluation.

Restructuring the process and the materials (S) so that applications are evaluated per criterion instead of applicant reduces the likelihood of selectively downplaying or inflating applicants' strengths and weaknesses. Similarly, asking evaluators to emphasize inclusion arguments (why they think an application should be evaluated positively) circumvents (always) giving the benefit of the doubt to some and (only) raising doubts about others. Having the same panelists ask the same questions in the same order to all applicants, a basic recommendation to improve the predictive validity of job interviews is a practice not yet often carried out in panel interviews.

In addition, we stress that applicants, applications and scores are not to be discussed outside the meeting of the panel, to make sure that everyone involved has access to the same information and arguments, and to ensure accountability and transparency of the process. Evaluators should fill out the evaluation forms completely and individually before sharing with other evaluators, including an explanation of high and not only low scores.

Finally, time should be strictly kept (T), to ensure that equal time is spent on each criterion and application and that enough time is left for the calibration. While discussing the meaning of criteria and deciding on some ground rules for the panel discussion and the interviews takes time, it will also save time during the deliberation and calibration because of a shared understanding of the criteria and the process.

Throughout our argumentation, in the vignettes and the quotes, the role of the panel chair has appeared as crucial in the evaluation and decision process. We underline this importance here, providing a few instances in which the chair can take responsibility in process optimization. In the preparation of the meeting, in instructing (new) panel members and in the calibration process at the end of the meeting, the chair can effectively make or break the effort. Some chairs take on the role of referee or assign the role of referee for someone to observe the process, take notes (using an app) and provide feedback on the occurrence of (linguistic) bias and opportunities for improving the process. An external observer (physically present or in online meetings), who is explicitly not a panel member but does have status and voice can be engaged for this role as well. We applaud chairs who encourage and incorporate both positive and negative feedback and who allow time for careful calibration and reconsideration at the end of a long and intense evaluation round (see [Table 4](#) for general recommendations and wild ideas).

Discussion

In this professional insights contribution, we have provided both specific and general recommendations for process optimization and discretion elimination in the evaluation and decision process for research funding. Our argumentation and recommendations serve to problematize the requirement of "(over)selling science" as an aspect adding additional, complex and gendered layers on top of the general challenge of making research funding more inclusive. In this sense, our recommendations can be translated to other performance evaluations in the editorial process or in research organizations. Especially in recruitment and selection, person-organization fit is an important criterion (relative to being able to sell science) that may lead to similar (dis)advantage for some compared to others depending on perceived (lack of) fit to an implicit ideal image ([Herschberg, 2019](#); [Rivera, 2017](#); [White-Lewis, 2020](#)).

When we design and implement interventions for research organizations based on these arguments, the question is often asked whether there is an optimum to process optimization in terms of standardization. Indeed, in some cases, algorithms (if unbiased), wild cards and

lotteries can take over (some of) the decision-making and reduce (some of) the time and other resources involved in instructing and supporting panels and panel chairs. However, not all funding schemes, evaluation criteria and process steps are identical, and these interventions work best when customized to the particular context. In this sense, customization and standardization go hand in hand in bias mitigation, process optimization and discretion elimination in research funding evaluation and decisions.

Note

1. Interview quotes are provided in the supplementary material submitted with this manuscript

References

- Bickmore, W., Cunningham-Burley, S. and Frame, M. (2018), "Funding mechanisms risk promoting conscious bias", *Nature*, Vol. 562 No. 7726, p. 192, doi: [10.1038/d41586-018-06979-w](https://doi.org/10.1038/d41586-018-06979-w).
- Biernat, M., Carnes, M., Filut, A. and Kaatz, A. (2020), "Gender, race, and grant reviews: translating and responding to research feedback", *Personality and Social Psychology Bulletin*, Vol. 46 No. 1, pp. 140-154, doi: [10.1177/0146167219845921](https://doi.org/10.1177/0146167219845921).
- Bleijenbergh, I.L., Van Engen, M.L. and Vinkenburgh, C.J. (2013), "Othering women: fluid images of the ideal academic", *Equality, Diversity and Inclusion*, Vol. 32 No. 1, pp. 22-35, doi: [10.1108/02610151311305597](https://doi.org/10.1108/02610151311305597).
- Bloch, C., Graversen, E.K. and Pedersen, H.S. (2014), "Competitive research grants and their impact on career performance", *Minerva*, Vol. 52 No. 1, pp. 77-96, doi: [10.1007/s11024-014-9247-0](https://doi.org/10.1007/s11024-014-9247-0).
- Bol, T., de Vaan, M. and van de Rijt, A. (2018), "The Matthew effect in science funding", *Proceedings of the National Academy of Sciences*, Vol. 115 No. 19, pp. 4887-4890, doi: [10.1073/pnas.1719557115](https://doi.org/10.1073/pnas.1719557115).
- Correll, S.J., Weisshaar, K.R., Wynn, A.T. and Wehner, J.D. (2020), "Inside the black box of organizational life: the gendered language of performance assessment", *American Sociological Review*, Vol. 85 No. 6, pp. 1022-1050, doi: [10.1177/0003122420962080](https://doi.org/10.1177/0003122420962080).
- Fox, M.F., Whittington, K.B. and Linkova, M. (2017), "Gender, (In)equity, and the scientific workforce", in Felt, U., Fouche, R., Miller, C.A. and Smith-Doerr, L. (Eds), *Handbook of Science and Technology Studies*, MIT Press, Cambridge, MA, p. 50.
- Heilman, M.E., Manzi, F. and Braun, S. (2015), "Presumed incompetent: perceived lack of fit and gender bias in recruitment and selection", in Broadbridge, A.M. and Fielden, S. (Eds), *Handbook of Gendered Careers in Management*, Edward Elgar, Cheltenham, pp. 90-104, doi: [10.4337/9781782547709](https://doi.org/10.4337/9781782547709).
- Herschberg, C. (2019), "Through the gate of the neoliberal academy, the (re) production of inequalities in the recruitment and selection of early-career researchers", *Radboud University*, PhD Thesis, available at: <https://repository.ubn.ru.nl/handle/2066/205633>.
- Kaatz, A., Magua, W., Zimmerman, D.R. and Carnes, M. (2015), "A quantitative linguistic analysis of national institutes of health R01 application critiques from investigators at one institution", *Academic Medicine*, Vol. 90 No. 1, pp. 69-75, doi: [10.1097/acm.0000000000000442](https://doi.org/10.1097/acm.0000000000000442).
- King, E.B., Avery, D.R., Hebl, M.R. and Cortina, J.M. (2018), "Systematic subjectivity: how subtle biases infect the scholarship review process", *Journal of Management*, Vol. 44 No. 3, pp. 843-853, doi: [10.1177/0149206317743553](https://doi.org/10.1177/0149206317743553).
- Leslie, S.-J., Cimpian, A., Meyer, M. and Freeland, E. (2015), "Expectations of brilliance underlie gender distributions across academic disciplines", *Science*, Vol. 347 No. 6219, pp. 262-265, doi: [10.1126/science.1261375](https://doi.org/10.1126/science.1261375).
- Madera, J.M., Hebl, M.R., Dial, H., Martin, R. and Valian, V. (2018), "Raising doubt in letters of recommendation for academia: gender differences and their impact", *Journal of Business and Psychology*, Vol. 34, pp. 287-303, doi: [10.1007/s10869-018-9541-1](https://doi.org/10.1007/s10869-018-9541-1).

- Malmström, M., Johansson, J. and Wincent, J. (2017), "Gender stereotypes and venture support decisions: how governmental venture capitalists socially construct entrepreneurs' potential", *Entrepreneurship Theory and Practice*, Vol. 41 No. 5, pp. 833-860, doi: [10.1111/etap.12275](https://doi.org/10.1111/etap.12275).
- Markowitz, D.M. (2019), "What words are worth: national science foundation grant abstracts indicate award funding", *Journal of Language and Social Psychology*, SAGE Publications, Vol. 38 No. 3, pp. 264-282, doi: [10.1177/0261927X18824859](https://doi.org/10.1177/0261927X18824859).
- Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T. and Huguet, P. (2019), "Committees with implicit biases promote fewer women when they do not believe gender bias exists", *Nature Human Behaviour*, Vol. 3, pp. 1171-1179, doi: [10.1038/s41562-019-0686-3](https://doi.org/10.1038/s41562-019-0686-3).
- Rivera, L.A. (2017), "When two bodies are (not) a problem: gender and relationship status discrimination in academic hiring", *American Sociological Review*, Vol. 82 No. 6, pp. 1111-1138, doi: [10.1177/0003122417739294](https://doi.org/10.1177/0003122417739294).
- van den Besselaar, P., Sandström, U. and Schiffbänker, H. (2018), "Studying grant decision-making: a linguistic analysis of review reports", *Scientometrics*, Vol. 117, pp. 313-329, doi: [10.1007/s11192-018-2848-x](https://doi.org/10.1007/s11192-018-2848-x).
- van Veelen, R. and Derks, B. (2020), "Academics as superheroes: female academics' lack of fit with the masculine stereotype of success limits their career advancement", *PsyArXiv*. doi: [10.31234/osf.io/c3k56](https://doi.org/10.31234/osf.io/c3k56).
- Vinkenburg, C.J. (2017), "Engaging gatekeepers, optimizing decision making, and mitigating bias: design specifications for systemic diversity interventions", *The Journal of Applied Behavioral Science*, Vol. 53 No. 2, pp. 212-234, doi: [10.1177/0021886317703292](https://doi.org/10.1177/0021886317703292).
- Vinkenburg, C.J., Jansen, P.G.W., Dries, N. and Pepermans, R. (2014), "Arena: a critical conceptual framework of top management selection", *Group and Organization Management*, Vol. 39 No. 1, pp. 33-68, doi: [10.1177/1059601113492846](https://doi.org/10.1177/1059601113492846).
- White-Lewis, D.K. (2020), "The facade of fit in faculty search processes", *The Journal of Higher Education*, Vol. 91 No. 6, pp. 833-857, doi: [10.1080/00221546.2020.1775058](https://doi.org/10.1080/00221546.2020.1775058).

Supplementary Material

The supplementary files are available online for this article

Corresponding author

Claartje J. Vinkenburg can be contacted at: c.j.vinkenburg@gmail.com

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com