

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:

2024-08-27

Deposited version:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Santos, J. M. (2024). Who gets the grant? A persona-based investigation into research funding panelist preferences. *Research Evaluation*. 33

Further information on publisher's website:

[10.1093/reseval/rvae030](https://doi.org/10.1093/reseval/rvae030)

Publisher's copyright statement:

This is the peer reviewed version of the following article: Santos, J. M. (2024). Who gets the grant? A persona-based investigation into research funding panelist preferences. *Research Evaluation*. 33, which has been published in final form at <https://dx.doi.org/10.1093/reseval/rvae030>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Who Gets the Grant? A Persona-Based Investigation into Research Funding Panelist Preferences

João M. Santos^a *

^a Iscte – Instituto Universitário de Lisboa, Center for Research and Studies in Sociology (Cies_Iscte), Lisbon, Portugal

* Corresponding author:

E-mail: joao_marques_santos@iscte-iul.pt

Address: Cidade Universitária de Lisboa, Avenida das Forças Armadas, 1649-029 Lisboa, Portugal

ORCID: <https://orcid.org/0000-0001-6414-6891>

Abstract

The allocation of scientific funding through grant programs is crucial for research advancement. While independent peer panels typically handle evaluations, their decisions can lean on personal preferences that go beyond the stated criteria, leading to inconsistencies and potential biases. Given these concerns, our study employs a novel method, using simulated, data-driven, and narrative personas of fictional candidates, to identify the attributes that significantly influence panelist choices and profile what an optimal candidate would look like. Our findings reveal a preference for mid-career, multidisciplinary researchers with significant publications, citations, and prior project experience. Such optimal candidates also lean toward applied science, collaborative research, interactions with both industry and the public, and a progressive stance on science. Contrary to existing literature, this study found indications of a bias toward female candidates, which we interpret as a result of deliberate correction caused by awareness of existing biases in academia. Age emerged as another influential factor, suggesting either a preference for younger researchers or a perception of waning productivity among seasoned academics. Consistency in a candidate's profile, rather than standout attributes, was also favored by panelists.

Keywords: research funding, panels, research evaluation, personas

Conflicts of Interest

The author declares that there exists no competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) [2020.03729.CEECIND].

1. Introduction

Funding is essential for the advancement of science, and as such, the processes through which funds are allocated are of vital importance. Grant programs typically employ panels of independent peers to evaluate submissions; these panels serve both as a quality control mechanism (Alberts et al., 2008) and aim to determine the most optimal allocation of funds in applications of similar merit. However, this approach has been noted in some cases as subjective, as panelists often must make choices under uncertainty or have differing interpretations of the evaluation guidelines (Zhu et al., 2022). Because of this, agreement among panelists tends to be low (Jerrim & Vries, 2020), with the approval of some applications being more dependent on the reviewers who were assigned than on the application itself (Pier et al., 2018). The fact that two different panelists can have opposite opinions regarding an application in the presence of explicit evaluation criteria suggests, then, that a substantial part of the decision-making depends on panelist preferences and not necessarily on the criteria. Previous research on the topic has suggested that affect might play a more important role in the choice than the contents of the application itself (Santos, 2022).

This notion was indeed a cornerstone of seminal works such as *How Professors Think* (Lamont, 2009), which introduced the notion of “evanescent criteria” - implicit and intrinsic criteria that are employed by panelists and that largely relate to the candidate’s attributes - such as perceived intelligence, personal features, and moral qualities. The degree to which these criteria influence the appraisal of the entire application is unknown but is likely to be significant; it has been suggested that certain features of an application can trigger a “halo effect” that can influence overall appraisal (Santos, 2022). This partially explains the persistence of certain biases that are based, for example, on reputation (Murray et al., 2016) or gender (Schiffbaenker et al., 2022). As these criteria are not based on scientific merit, it is not surprising that there has been a push to move toward lottery-based systems, which are less susceptible to such biases (Smaldino et al., 2019).

However, the literature also notes the existence of biases that are more scientific in nature. For example, it has been demonstrated that panelists tend to prefer researchers who are prolific (Arora et al., 1998), mono-disciplinary (Banal-Estañol et al., 2019), and conservative (Ayoubi et al., 2021). These are a combination of attributes that do not necessarily suggest a trailblazing profile for winning candidates. In fact, it has been noted that the success of a proposal is not related to its quality but, rather, to the applicant’s aptitude for securing funding (Laudel, 2006), a mechanism that has led to the concentration of funding among elites (Larivière et al., 2010), an approach that tends to lead to diminishing returns (Mongeon et al., 2016). Nevertheless, it is important to note that, at least recently, panelists have become aware of these mechanisms, and recent research notes several shifts in preferences that can be perceived as attempts to correct long-standing biases (Bol et al., 2022; Lawson & Salter, 2023). This is a balancing act, as there is always the possibility of overcorrecting a bias.

Given the above, it is evident that an applicant’s understanding of the preferences of panelists is perhaps more important than an understanding of the rules of the grant program itself. Although applications typically contain two aspects - the candidate’s profile and a project proposal - given the aforementioned context and the importance of individual attributes that are well-grounded in the literature, this exercise focuses exclusively on the participant’s characteristics and does not tackle the project facet, which is an altogether different challenge. Although literature on the topic is emerging, there is still a shortage of quantitative studies on panelist choices, especially concerning candidate profiles. This is unsurprising given that this data is typically not disclosed by funding agencies. In this study, we aim to tackle this challenge by employing a novel approach: we present panelists with simulated data-driven and narrative personas representing fictional candidates and

seek to determine the attributes that influence their choices the most, thereby determining “who gets the grant.” As a result, we hope to offer insights that could guide both future applicants and panels, potentially reshaping the way in which grant allocations are approached.

2. Method

2.1. Data collection

To collect our data, we contacted 20,807 academics who were registered as corresponding authors in Scopus-indexed articles, based on a random search of articles across all fields of science. These academics were contacted by email during June and July of 2023 with an invitation to participate in this study. Before participants were able to proceed with the survey itself, they were required to read and agree to an informed consent form that was presented online. In total, 3,462 participants began the process, a response rate of 17%. Immediately after completing the consent form, participants were asked if they had ever served on a research grant evaluation panel. In this exercise, we employed only the data for individuals who reported having served as panelists. As such, we filtered out 684 participants who reported not having served as panelists. Finally, among the panelists, we filtered out 470 participants who did not finish the survey, as due to the nature of this exercise, the data was largely not imputable. This led to a final sample of 1,882 participants. A total of 68.11% (N = 1,282) of the participants were males and 31.88% (N = 600) were females; on average, participants were 58.21 years old (SD = 11.33). In terms of discipline, 3.72% (N = 70) were from the Agricultural Sciences, 18.23% (N = 343) from Engineering and Technology, 3.13% (N = 59) from the Humanities, 24.65% (N = 464) from the Medical and Health Sciences, 25.03% (N = 471) from the Natural Sciences, and 25.24% (N = 475) from the Social Sciences. The sample included participants from numerous countries, of which the most represented were the United States (N = 394; 20.94%), Italy (N = 130; 6.91%), Spain (N = 119; 6.32%), the United Kingdom (N = 119; 6.32%), and Canada (N = 97; 5.15%).

2.2. Design

As simulating an entire panel, including fictional projects and an entire pool of candidates at once, would render the entire exercise too complex and time-consuming for the participants, we presented participants with a deliberately minimalist simulation. To preserve the competitive nature of a panel, the fictional candidates could not be presented in a vacuum; they had to be introduced in contrast with other participants. To solve this, we opted to conduct this study as a Conjoint Analysis exercise (Rao, 2014) and present the candidates in pairs, with the participant asked to choose either one or the other. Conjoint Analysis is a methodology often used in marketing studies in which a series of products or services with controlled sets of varying attributes are presented to respondents for their choice. This makes it possible to identify the respondent’s preferences regarding the choice’s attributes; as such, this methodology was ideal for the purposes of the study.

The fictional grant was framed as a national-level grant program aiming to support projects for the general advancement of knowledge, encompassing both basic and applied science. Participants were informed that the project evaluation stage had already taken place and that they were assigned the role of breaking ties between candidates with similar scores. They were also informed that the regulations did not specify any criteria for the tiebreak and that the panel had agreed to defer to the participant’s best judgment. Finally, the participants were informed that they should assume that all presented candidates had the exact same evaluation in their projects and that their skills were perfectly aligned with their proposal; thus, the decision should derive solely from their traits, accolades, and potential.

The exercise, as it was presented to the participants, comprised two rounds. In one round (henceforth, “*Quantitative Personas*”), we presented quantitative, data-driven, and procedurally generated personas (Jansen et al., 2022) with a limited amount of information, notably, age, gender, publications (including solo-authored), citations, international experience, and project experience. These attributes were chosen as some that are known to influence panelist preferences based on previous literature (e.g., Arora et al., 1998; Santos, 2022; Seeber et al., 2022). Each participant was presented with five pairs of personas and was asked to choose one from each pair.

In the other round (henceforth, “*Narrative Personas*”), we presented narrative text-based personas (e.g., Salminen et al., 2020), including a short biography, a name, and their most distinguishing features. These personas each exhibited one salient feature and were presented with a contrasting counterpart, for example, a generalist versus a specialist. For this round, participants were presented with seven pairs of personas and were asked to choose one from each. After selecting their seven candidates, they were presented with one final choice: they were presented with all of their previously chosen candidates and were invited to select one as the winner of this fictional grant program.

To avoid order effects, the quantitative and narrative comparisons were randomized so that either one could be presented first or second. In the next sections, we describe the generation process for the personas employed in each exercise.

2.3. Quantitative Persona Generation

First, one highly important feature of the generation of quantitative personas that is key to the research design of this study: as we aimed to use the generated personas as “cases” in our analysis, we needed to ensure a wide variety of values, similar to what would happen in a sample with real individuals. Because of this, each generated persona needed to be unique for each participant. To achieve this, we developed a lightweight procedural persona generator in JavaScript and embedded it into the surveying platform (Qualtrics). As a result, each participant was presented with entirely unique personas. Using this approach meant that each participant yielded 10 cases for the analysis.

Second, to make the choices meaningful and not one-sided, it was necessary to ensure that both presented personas were balanced, each with their own strengths and weaknesses. To facilitate this, instead of generating two separate personas, for each choice we generated a “seed” persona and then created a derivative persona based on the seed persona. The seed persona was data-driven and generated based on real-world data (notably, age ranges, publications, and citations).¹ Thus, we began by generating the seed persona following these steps:

1. First we generated *age*, a random number between 45 and 65 years.
2. Then, we determined the personas’ *publications* and *citations*. Using our available real-world data, we computed 20 percentile ranges, normalized by age (i.e., publications and citations per year of age). One of these percentile ranges was randomly selected, and a random number was chosen within the lower and upper bounds of that percentile range. The values were multiplied by the age generated in Step 1.
3. *Gender* was randomized to either male or female (with a 50% chance of each).
4. The *solo paper percentage* was generated by randomly selecting a number between 0.1 and 0.5. An absolute value was generated at the end of the process by multiplying this percentage with the generated publication count.

¹ For this purpose, we employed a large dataset with the bibliometric data of 12,169 researchers which was generated as part of a previous project.

5. *For international jobs held and project experience*, a random number was generated between 0 and 4, inclusive. This represents a Likert scale; each number was qualitatively presented with a corresponding label (for example, 2 in international jobs was presented as “the candidate has held some international positions”; 4 in projects experience was presented as “the candidate has participated in a substantial number of research projects”).

Now that the seed persona had been generated, we could create a derivative persona. The gender was always set to be the opposite of the seed persona. Then, we took the seed persona and randomly selected three attributes to be enhanced and three attributes to be penalized. This randomization was done using the Fisher–Yates shuffle to ensure that over a sufficiently large N, each attribute was selected for enhancement and for penalization an equal number of times. For this purpose, we considered that enhancing *age* translated into reducing its numeric value, and all other values were increased. Table 1 indicates the bonuses and maluses and their value range. All are multiplicative of the seed value except project experience and international job experience, which are additive.

Table 1: Bonuses and maluses for generating the derivative persona

	Malus	Bonus
Age	* [1.20 – 1.30]	* [0.70 – 0.80]
Publications	* [0.50 – 0.75]	* [1.25 – 1.50]
Citations	* [0.50 – 0.75]	* [1.25 – 1.50]
Solo Paper Percentage	* [0.50 – 0.75]	* [1.25 – 1.50]
Projects Experience	-2	+2
International Job Experience	-2	+2

Due to this multiplicative nature, a mechanism was implemented to handle out-of-bounds values (e.g., very high ages and projects or international scores going beyond the 5-point Likert scale). *Age* was clamped to a lower bound of 35 and an upper bound of 80; *projects* and *international job experience* were clamped to values between 0 and 4, and *solo paper percentage* was clamped to values between 0.10 and 0.75. Whenever the generated value was out-of-bounds, we set it to either the minimum or maximum value, depending on the direction of the overflow, and computed the difference between the original value and the clamped value as an offset. Then, we applied this offset to the seed persona as well, to preserve relative balance.

As the reader might have noticed, the seed persona’s age and solo paper percentages were set to be lower than the derivative’s clamped values. This was to minimize the number of times that out-of-bounds values were generated due to the multiplicative nature of the bonuses and maluses. A drawback of this generative process, however, is that the numeric range of values for the derivative persona was higher than for the seed persona. Finally, to avoid any possibility of an order effect, we randomized the positioning of the seed and derivative personas so that each had a 50% chance of being presented as either Candidate 1 or Candidate 2. Table 2 presents the mean and standard deviation of a simulation of 100,000 personas and compares them with those that were actually generated during the study. As can be observed, the process worked as intended, and the means between the seed and derivative personas are extremely close, ensuring that they are as numerically balanced as possible despite the fact that participants were presented with extremely different candidates.

Table 2: Descriptives for seed and derivative personas, simulated and deployed

	Simulated				Deployed			
	Seed		Derivative		Seed		Derivative	
	(N = 50,000)		(N = 50,000)		(N = 13,938)		(N = 13,938)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Gender	0.500	0.499	0.499	0.499	0.502	0.502	0.497	0.500
Age	55.039	5.885	56.76	15.026	55.304	8.186	55.272	15.195
Publications	56.761	39.056	56.726	47.082	57.795	40.641	57.265	48.142
Citations	1365.098	1091.675	1360.152	1273.806	1387.781	1129.039	1388.913	1322.887
Solo Paper %	17.180	14.168	16.830	18.062	17.447	14.599	16.766	17.612
Projects Exp.	2.000	0.998	2.000	1.613	2.003	0.998	2.023	1.611
International Exp.	1.999	1.004	2.002	1.609	2.009	1.007	2.011	1.610

2.4. Narrative Persona Generation

To generate the narrative personas, several short descriptions of fictional researchers were written. These personas are based on archetypes which are grounded in ongoing debates and literature (further expanded in greater depth in the analysis section, further ahead), and are aimed to represent contrasting facets that can be displayed by researchers within academia, for example, “The Generalist” or “The Specialist”. The descriptions were standardized to ensure balance - both personas were set to be of the same gender and of similar age. No outstanding features beyond the one that was of interest were included in the description, but some minor “flavor text” was included to lend credibility to the persona while maintaining balance and focusing on the feature of interest.

The descriptions were always presented following the same structure: an initial paragraph with a mini-bio, followed by a shorter paragraph emphasizing the feature or concept that was represented by the persona. After the description, the question itself provided a brief summary of the two candidates to remind the participants of the candidates being chosen.

The choices were presented in a randomized order. After the seven choices, participants were asked to make one final selection for the fictional grant winner and were presented with the seven candidates they had previously chosen. Again, for the purposes of reminding participants of their profiles without requiring them to re-read the entire bios again, short one-line descriptions were provided. The various persona pairs, as well as their names, archetype names, and descriptions, are provided in Table 3. The supplementary materials of this article contain the full text description of each persona as presented to the participants. For the ease of reading this article without the reader having to memorize the persona names, only the archetypes’ names will be used in the subsequent sections.

Table 3: Description of the narrative personas

Pair	Persona Name	Archetype	Short Description
1	Dr. John Smith	Theorist	A revered theorist known for his deep intellectual depth and significant theoretical contributions to his field.
	Dr. Alex Carter	Technologist	An innovative trailblazer who consistently integrates pioneering methodologies and cutting-edge technologies into his research.
2	Dr. Leonard Fischer	Fundraiser	A successful grant-seeker who sustains and grows research through consistently securing funding for various projects.

	Dr. Tom Edwards	Underdog	A patient academic with a slower pace and little success so far, but with a high-potential project.
3	Dr. Angela Martin	Networker	A master networker known for her ability to foster collaborations and drive joint research projects across institutions.
	Dr. Jane Wallace	Soloist	A self-motivated independent researcher with a remarkable collection of solo-authored works in her field.
4	Dr. Laura Wilson	Public Intellectual	A public intellectual renowned for making complex academic ideas accessible and engaging to the general public.
	Dr. Samantha Smith	Pracademic	A successful pracademic who expertly bridges academia and industry, applying theoretical concepts to practical real-world scenarios.
5	Dr. Edward Davis	Traditionalist	A staunch traditionalist dedicated to preserving and championing the established scientific system.
	Dr. Jasper Sullivan	Iconoclast	An advocate for scientific evolution with innovative ideas for a dynamic, flexible system that fosters progress.
6	Dr. Mary Sullivan	Elite Educator	A passionate educator committed to teaching and mentoring the next generation of scholars, while also contributing to her academic field.
	Dr. Jennifer Wolf	Prolific Researcher	A highly productive researcher with substantial academic contributions in a short span of time, showing potential for even greater future achievements.
7	Dr. Oliver Harper	Generalist	A dynamic scholar celebrated for his broad, interdisciplinary approach to research and his ability to weave insights from various academic fields.
	Dr. Aaron Mitchel	Specialist	An esteemed academic with a passion for in-depth, specialized research, and a proven record of advancing his niche field.

Note: the archetype name was only used internally and never shown to the participants.

Most of the contrasting facets should be self-explanatory but are described later in the results section.

2.5. Procedure

The quantitative persona analysis makes use of a multilevel mixed-effects logistic regression. To accomplish this, the data set was transposed from the wide format into the long format, with each persona representing a case. A *Chosen* dummy variable was created for each persona, assuming a value of 1 if the persona was chosen and 0 otherwise. This was our dependent variable. Additionally, we created unique identifiers for each pair and for each participant to account for the hierarchical nature of the data. As such, in our multilevel regression, personas were nested within pairs, which were nested within participants. We also employed a TwoStep Cluster Analysis (Zhang et al., 1996) in a second step group participants into probability clusters, in order to subsequently identify the characteristics of the persona most likely to be chosen.

We included several control variables in our model. The *field of science (FOS)* controlled for the disciplinary area of the participant, as evaluation criteria are certain to differ by field; we employed Agricultural Sciences as a reference category. For the same purpose, we included the participant's *Continent* as a control variable to account for evaluation aspects which might differ by geographical location – although the country would be a more granular option, the numerous levels of such a

variable render it unsuitable for this type of model. *Participant gender* was a dummy controlling for participant gender, using males as a reference category, while *participant age* was a continuous variable representing and controlling for the participant's age. Finally, *persona type* was a dummy that controls for whether the persona was a seed or a derivative persona, which are described above. Despite the steps taken to ensure the numerical balance of the personas, derivatives were found to be chosen much more often than seed personas; thus, it became necessary to control for this effect through a fixed factor.

Finally, the variables of interest were the attributes of each persona. *Persona gender* indicated the persona's presented gender (male as a reference category), and *persona age* was the age that was presented for the persona. *Persona publications*, *solo publications*, and *citations* all referred to the bibliometric data that was presented for the persona, while *persona international experience* and *persona project experience* represented the numeric value for the qualitative label that was generated for the persona (expressed on a 0–4 scale). Finally, for the same reason that led to the inclusion of *persona type*, a measure of *Persona Variability* was also introduced. This variable results of the computed variance for a given persona's standardized (through z-scores) numerical attributes and represents a measure of cohesion or dispersion in the generated attributes.

For the narrative persona analysis, we employed a more descriptive approach and aimed to characterize the most-chosen attribute and contextualize it within the existing literature to make sense of the findings. As such, this section is more qualitative in nature.

3. Results

3.1. Quantitative Persona Analysis

This analysis begins by conducting a multilevel mixed effects logistic regression, nesting observations (individual personas) within the pairs in which they were presented and nesting these pairs within the individual participants, also controlling for a series of fixed effects. The resulting model exhibited a predictive accuracy of 62.16%, with a sensitivity of 65.83% and specificity of 58.54%, which is appropriate given the exploratory nature of this study. The results of this analysis are summarized in Table 4:

Table 4: Multilevel mixed effects logistic regression.

Variable	Odds ratio	Std. Err.	95% Confidence Interval	
			Lower Bound	Upper Bound
FOS (Engineering and Technology)	0.968	0.024	0.921	1.016
FOS (Humanities)	0.982	0.032	0.920	1.048
FOS (Medical and Health Sciences)	0.958	0.023	0.914	1.005
FOS (Natural Sciences)	0.954	0.023	0.090	1.001
FOS (Social Sciences)	0.973	0.024	0.927	1.021
Continent (Africa)	1.021	0.037	0.951	1.096
Continent (Asia)	0.995	0.018	0.959	1.033
Continent (Europe)	1.007	0.010	0.986	1.028
Continent (Oceania)	1.021	0.023	0.976	1.068
Continent (South America)	0.979	0.019	0.941	1.018
Participant Gender (Female)	1.006	0.009	0.987	1.026
Participant Age	0.999	0.000	0.999	1.000
Persona Type (Derivative)	0.566 ***	0.027	0.514	0.623
Persona Gender (Female)	1.978 ***	0.079	1.828	2.141
Persona Age	0.960 ***	0.001	0.958	0.963
Persona Publications	1.003 ***	0.000	1.002	1.004

Persona Solo Publications	0.995 ***	0.001	0.993	0.997
Persona Citations	1.000 *	0.000	1.000	1.000
Persona International Experience	0.938 ***	0.012	0.915	0.963
Persona Project Experience	1.144 ***	0.015	1.114	1.174
Persona Variability	0.843 ***	0.027	0.792	0.898
Constant	7.559 ***	0.678	6.000	8.676

Observations 18,820

Number of groups 1,882

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Odds ratios are shown, with robust standard errors and 95% confidence interval. Observations are nested within pairs, which are nested within participants. Reference category for FOS is Agricultural Sciences; reference category for Continent is North America.

The first notable finding is that, despite the multiple measures that were taken to ensure the balance of the two personas - which, as indicated above, had similar means across all attributes - derivative personas had roughly 50% less probability of being chosen ($OR = -0.566$, $p < .001$). As the order of presentation was shuffled, thus precluding any sort of order effect, and their means were identical over a sufficiently large number of generations, only two possible explanations remained for this unexpected finding: that this could be caused by the relatively larger range of values for derivative personas or that it could be caused by a higher dispersion of values, both of which are unavoidable byproducts of the random component of the generation process. We re-estimated a diagnostic model using only derivative personas with generated values within the same maximum and minimum range as the ones generated for seed personas, and the result was still consistent ($OR = 0.722$, $p < .001$). Thus, the only remaining explanation is the larger standard deviation, which is corroborated by the effect of Persona Variability, also a significant predictor ($OR = 0.843$, $p < .001$), with personas with a higher dispersion of attributes being less likely to be chosen. A possible interpretation is that participants preferred the personas that had more stable attributes across the board, as the generation process emphasized some attributes while penalizing others for the derivative persona. This is a surprising but serendipitous finding, as it suggests that consistency, at least in these presented traits, has an important role.

Age has a negative impact on the odds of being chosen: each additional year represents a 4% decrease in the probability of being selected ($OR = 0.960$, $p < .001$), indicating a preference toward younger candidates. Publications also demonstrate a significant effect but one that is not highly relevant: each additional paper represents only an additional 0.3% probability increase of being chosen ($OR = 1.003$, $p < .001$), whereas citations exhibit a effect that is significant but indistinguishable from zero ($OR = 0.000$, $p < .05$). Furthermore, solo-authored papers diminish the probability of being chosen; each such paper reduces this probability by 0.5% ($OR = 0.995$, $p < .001$). Considering the positive effect of a paper in tandem with the negative effect of a solo paper, publishing a single-authored article would, in total, result in a net loss in terms of probability of being selected.

International experience, as indicated by positions held abroad, also led to a lower probability of being selected: each level increase leads to a decrease of 6.2% in the probability of being chosen ($OR = 0.938$, $p < .001$). By contrast, past project experience has the opposite effect, with each level increase resulting in an increment of 14.4% in the probability of being chosen ($OR = 1.144$, $p < .001$).

Finally, we have the effect of the persona's gender, which merits special attention. All other variables being controlled for, female personas were 97.8% more likely to be selected than male personas ($O = 1.978$, $p < .001$). This finding is discussed in depth in the Discussion section, but it is

important to immediately note that we do not interpret this finding as indicative that panelists have a gender bias favoring women, especially given the abundance of literature and empirical data revealing that there is a bias but in the opposite direction. Three more likely interpretations, within the context of existing literature and data, are that this finding is due to bias awareness and overcompensation, perceived scrutiny, stereotype reversal, or a combination of all three. As mentioned, this topic is revisited later in the manuscript.

The next step in this analysis was estimating the probability of being selected using the previous model. This value was then used for a Cluster Analysis. The Cluster Analysis used the probability of being selected as the sole input variable; this was a deliberate choice as this variable already incorporates information from the predictor variables employed in the model (since they were used to compute the probability). The TwoStep Cluster Analysis algorithm identified three clusters with a silhouette measure of cohesion and separation of 0.7, which qualitatively indicates good fit for such a structure (Kaufman & Rousseeuw, 2009). After identifying the three clusters, we computed descriptive statistics for each of them to create profiles of personas with varying degrees of probability of being chosen. This is presented in Table 5, with Figure 1 showing the distribution of probabilities by cluster class:

Table 5: Cluster analysis using selection probability as input (other variables merely descriptive)

	Cluster 1	Cluster 2	Cluster 3
Probability	0.68	0.51	0.29
Cluster descriptives			
Gender	Female (90.2%)	Male (64.6%)	Male (70.8%)
Age	48.27	52.77	67.05
Publications	56.93	55.72	61.23
Solo Publications	15.90	16.21	20.00
Citations	1365.70	1343.53	1495.25
International Experience	1.70	1.96	2.45
Project Experience	2.30	1.97	1.76
Variability	0.60	0.71	1.31
N	7,588 (29.8%)	10,982 (43.1%)	6,932 (27.2%)

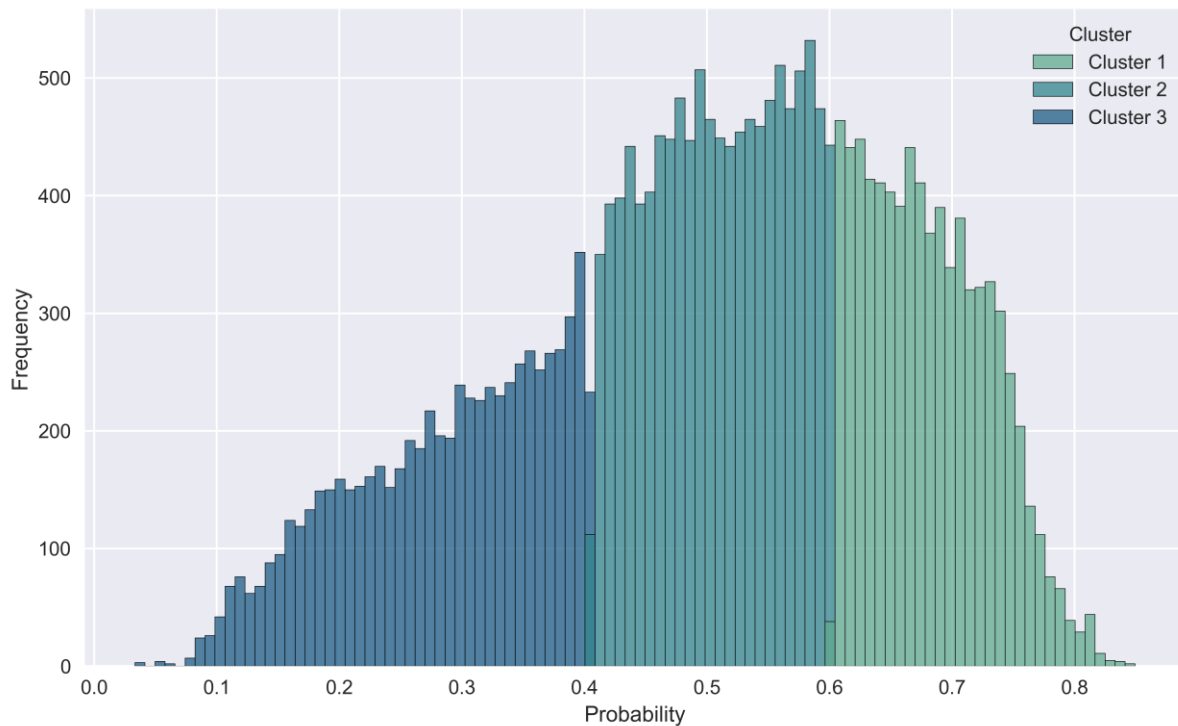


Figure 1. Distribution of probabilities by cluster class.

As the analysis reveals, the most distinguishing feature among the various clusters are the persona's gender, age, and attribute variability, with international experience and project experience also playing a role. Bibliometric indicators exhibit only modest differences across the various profiles. Thus, the persona profile with the highest probability of being selected (68%) represents a 48-year-old female who has published 57 articles, of which 16 were solo-authored, and has accrued 1,365 citations, having a balanced profile with modest variability. This persona archetype also has held a few international positions and has some experience with research projects. On the opposite end of the spectrum, the profile of the persona least likely to be selected (29%) is that of a 67-year-old male with 61 published articles, of which 20 were solo-authored, and having accrued 1,495 citations. This persona archetype has held many positions abroad, and has participated in a few research projects, and has an unbalanced array of attributes, exhibiting high variability. The middle-ground cluster (Cluster 2) neither leans towards being chosen or not (51% probability) – it essentially has the same bibliometric profile as Cluster 1, with the differences lying in being a slightly older (53 years) male, with slightly more international experience, slightly less project experience, and slightly more attribute variability.

3.2. Narrative Persona Analysis

The narrative personas were presented in a pairwise manner, as previously discussed, and represented contrasting features embodying existing tensions in academia. Figure 2 illustrates the percentage of times each persona was chosen within the pair. As the choices were presented in a random order, this section follows the same order that is indicated in the figure, which is sorted for ease of readability.

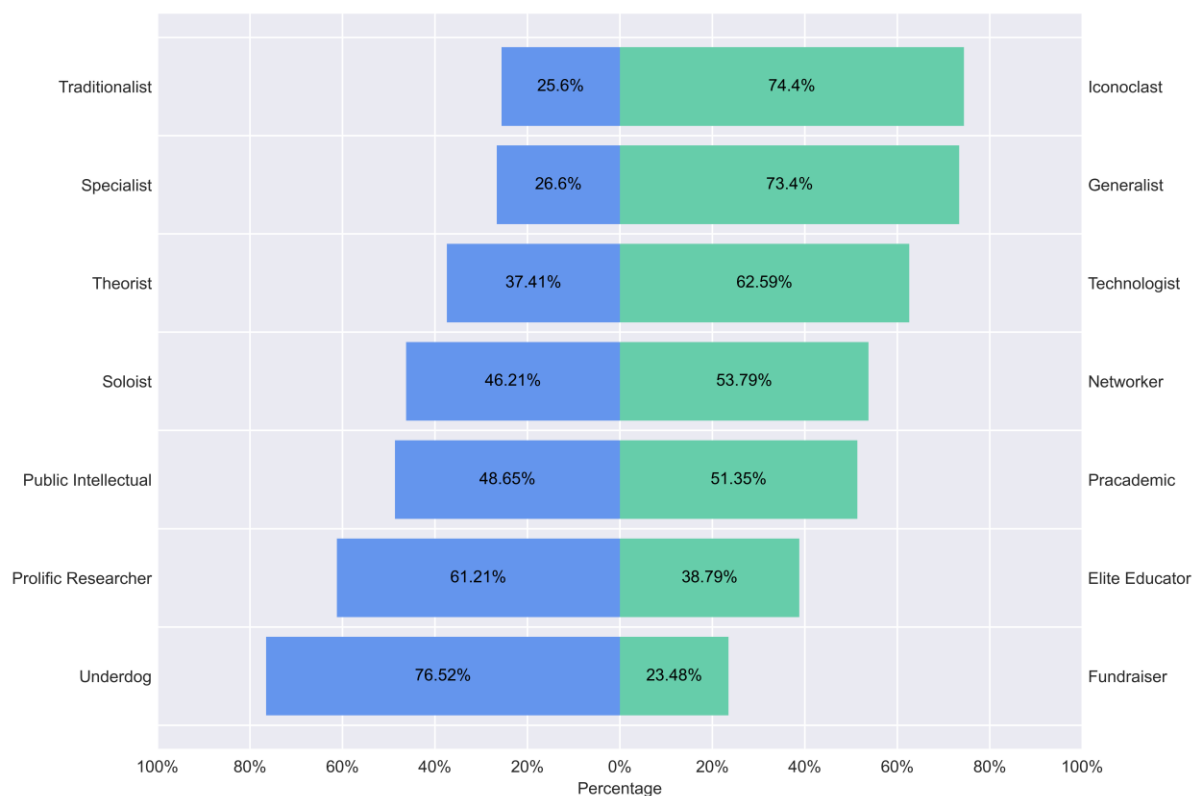


Figure 2. Persona preferences for each pair.

Traditionalist versus Iconoclast. This choice contrasts traditionalism in scientific practices (the Traditionalist) versus the pursuit of new scientific practices even when they challenge established norms (the Iconoclast). This tension can occur whenever changes emerge that demand a re-evaluation of existing practices, such as in the case of “normal science” versus “extraordinary science” (Kuhn, 2012), but it can also be seen in modern-day science as some established practices are criticized, as in the case of performance evaluation strategies overly relying on bibliometrics (Martin, 2011) or the systems currently used for the evaluation of applications (Smaldino et al., 2019). In this comparison, the Iconoclast was chosen 74.4% of the time when compared to the Traditionalist, who was selected 25.6% of the time. This suggests that panelists shy away from candidates who are more traditional in regard to scientific practices and prefer candidates who are outspoken on these topics, even when they challenge established norms.

Specialist versus Generalist. In this choice, we present two different ways to tackle scientific problems - specializing in a singular topic or branching out into a variety of disciplinary areas, contrasting depth of knowledge with breadth of knowledge. Both are valid approaches, but many problems in modern-day science favor a multidisciplinary approach due to their inherent complexity, which cannot be tackled with a mono-disciplinary approach (Martimianakis & Muzzin, 2015). Indeed, the pursuit of multidisciplinary endeavors has been demonstrated to be more beneficial at a professional level (Millar, 2013). The panelists’ preferences are aligned with this; 73.4% preferred the Generalist candidate when compared to the Specialist candidate.

Theorist versus Technologist. This choice represents the classic tension between basic science, represented by the Theorist, versus applied science, represented by the Technologist (Becher & Trowler, 2001). Both basic and applied science are required for progress, but basic science has one key feature: very often, its findings do not have immediate applications, which is not the case for applied science. Thus, basic science tends to be more reliant on public funding (Conceição et al.,

2004), and the private sector is often dependent on past public investment in basic science to pursue applied science (Mazzucato, 2013). Nevertheless, panelists exhibited a strong preference for candidates with a more applied profile, with the Technologist being chosen 62.59% of the time versus the Theorist, a candidate portrayed as being more focused on basic science, being selected 37.41% of the time.

Soloist versus Networker. In this choice, we contrast the representation of a traditional “ivory tower”-style researcher, who prefers to work on his or her own, with the representation of a researcher more engaged in collaborative ventures than in solo work. It is important to note that in some disciplinary fields, solo work is simply not an option; such is the case for many “hard” sciences in which large collaborative ventures are the norm, at least when compared to the “soft” sciences (Lauto & Valentin, 2013). Even so, a bibliometric study in the social sciences has noted that solo publishing tends to be associated with a lower quantity of publications, average journal impact factor, and overall impact of the published content (Haslam & Laham, 2009). Furthermore, researchers who reported having frequent invitations to collaborate have been demonstrated to accrue more publications overall and, as such, more citations (Santos et al., 2022). By contrast, it has also been demonstrated that solo articles are more rewarded in professional terms (notably, at the salary level; Hilmer & Hilmer, 2005), and publishing exclusively in collaborative efforts may be perceived as a lack of individual creativity and autonomy (Wray, 2002). For this reason, there are clear arguments and incentives for both pursuing solo work and collaborative work. As such, the results of this choice are not surprising, with a relatively even split between the Soloist, chosen 46.21% of the time, and the Networker, with a slight edge, having been chosen 53.79% of the time. This is aligned with the small net loss in probability of being selected for each published solo paper, which was noted in the previous analysis.

Public Intellectual versus Pracademic. This choice contrasts a researcher who specializes in public engagement with a researcher who applies knowledge within the industry. Both facets are considered desirable and they are not necessarily exclusive; however, it is relatively uncommon to find an individual who is engaged in both scientific communication and industrial applications. On the one hand, public engagement has been growing in importance and is indeed a common component in current research evaluation schemes (Reed et al., 2018; Rowe & Frewer, 2004), despite the fact that many researchers do not feel well-equipped to conduct public communication (Hamlyn et al., 2015). On the other hand, industry collaboration is also highly encouraged in research evaluation frameworks (Ma & Agnew, 2022), but despite this, direct engagement with the industry has not been revealed to have any impact on the relative success of grant applications (Öcalan-Özel & Llerena, 2021). Thus, the results of this choice are reflective of these tensions; again, there was a nearly even split between the Public Intellectual, who was chosen 48.65% of the time, and the Pracademic, who was chosen 51.35% of the time.

Prolific Researcher versus Elite Educator. This choice aims to reflect the tensions arising from the research-teaching nexus by contrasting a highly prolific academic who is heavily invested in research activities but lacks meaningful teaching experience with an academic who is highly experienced in teaching and has supervised numerous students but has made only a modest contribution to science in terms of publications. Although both activities traditionally have been seen as intertwined (Tight, 2016), a meta-analysis of various studies on the topic has revealed that they are, in fact, largely antagonistic activities (Marsh & Hattie, 2002). It has been suggested that a heavier teaching load can result in lower success rates in securing grants, as it leads to a relatively poorer publishing track record (Murray et al., 2016), a difficulty that is also noted by the academics themselves (Weenink et al., 2023). This suggestion is unsurprising given that teaching is seldom a criterion used for

evaluating grant applications. Because of this, the results are not surprising: 61.21% of the participants would choose the Prolific Researcher rather than the Elite Educator, who was selected 38.79% of the time.

Underdog versus Fundraiser. This choice has one distinguishing feature: it is the only one in which the candidates' projects are mentioned. The Fundraiser is portrayed as a seasoned researcher with an excellent track record of securing grants but who does not publish much and has a reliable but not groundbreaking project. Conversely, the Underdog is portrayed as a younger researcher, somewhat of an underachiever, who applied with a remarkable project with very high potential. A key feature of this choice is risk tolerance: the Fundraiser is clearly the "safer" choice (low risk, low reward), while the Underdog can be seen as a "riskier" choice (high risk, high reward). Additionally, the choice represents an existing issue in grant attribution: resources have historically been concentrated in candidates with a "winner" profile (Larivière et al., 2010), creating scientific elites; however, a concentration of resources in this manner has been demonstrated to lead to diminishing returns in innovation (Mongeon et al., 2016), making this strategy counterproductive for the advancement of science. The results of this choice reveal some awareness of this fact: the Underdog was chosen 76.52% of the time, versus 23.48% for the Fundraiser reflecting a preference for riskier candidates but also for spreading resources instead of concentrating them, a different and emerging approach that has also been noted in recent literature (Lawson & Salter, 2023).

Following the pairwise comparisons, participants were then presented with their previous choices and asked to select the final winner. The results of this final round are summarized in Figure 3.

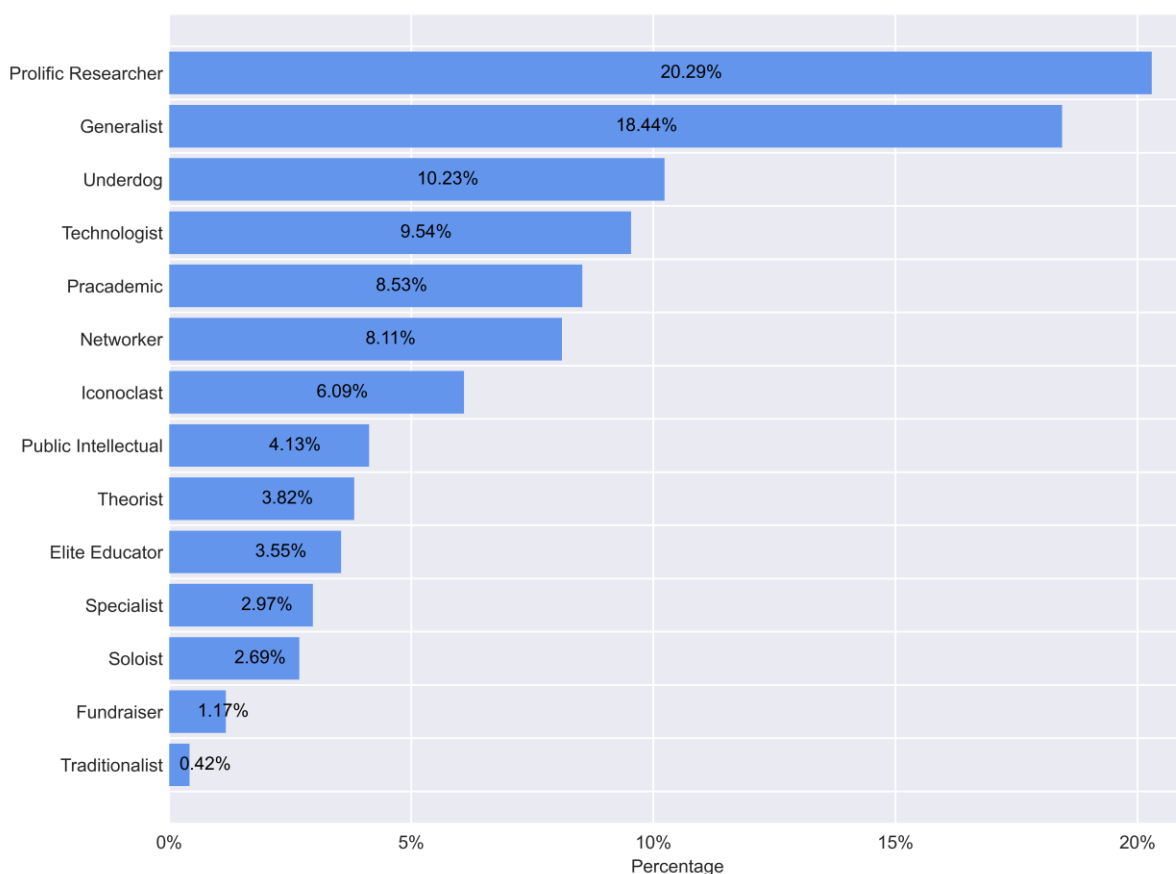


Figure 3. Global persona preferences.

This round is distinct from the previous ones because contrasting facets were no longer being presented but, rather, academics with a variety of attributes. As such, it aimed to identify the most desirable traits. Based on this, two stand at the forefront: the Prolific Researcher and the Generalist. The Prolific Researcher was chosen as the final winner 20.29% of the time, suggesting that grant applicants need to have a research-intensive profile to increase their chances of success. Furthermore, a multidisciplinary profile is also highly desirable, as the Generalist took second place, being chosen 18.44% of the time. The Underdog surprisingly was third but still ranked substantially lower than the aforementioned options, having been chosen 10.23% of the time. Readers will recall that this is the only participant for whom references were made to the project, which was described as outstanding. Despite the fact that the Underdog won against the Fundraiser (who was described as having a more conservative project) by the largest margin among all the pairwise comparisons, when the Underdog was pitted against candidates with more solid profiles, this persona was chosen much less often than the Prolific Researcher and the Generalist. Nevertheless, the Underdog ranked higher than the remaining candidates, with the closer personas being the Technologist, the Pracademic, and the Networker. This highlights the importance of a commendable project in a grant application, as this fictional candidate with a weaker profile was still able to rank relatively high in the comparison. The weaker-placed candidates were the Fundraiser and the Traditionalist; this suggests that panelists tend to avoid candidates with conservative profiles, as is the case of the Fundraiser (who specializes in securing grants but does not innovate much) and the Traditionalist (who actively defends the status quo in science).

4. Discussion and Conclusions

In this exercise, we aimed to characterize panelist preferences regarding candidates to profile what an optimal candidate would look like. Based on the cumulative findings of both the quantitative and the narrative persona analyses, this optimal candidate is a mid-career prolific researcher with a respectable number of published articles among which a notable number - roughly a quarter - were solo-authored, a reasonable number of citations received, and previous experience with research projects. In terms of approaches to science, this candidate is multidisciplinary, focuses on applied science, and embraces collaborative endeavors. The candidate has a mix of public outreach skills but also maintains connections to the industry. Additionally, the candidate has high-potential projects and is not afraid of challenging established norms. In this section, we examine some of the key findings of our analyses in further depth.

The first issue that needs to be addressed is the finding of the quantitative analysis, which suggests, at first glance, that panelists are biased toward female candidates. As mentioned in the results section, we do not directly interpret this as an indication of such a bias, given the overwhelming amount of empirical data that indicates there is a gender bias but one that favors men, not only in science in general (e.g., Abramo et al., 2016; Angervall, 2018; Santos et al., 2020) but also very specifically in grant evaluation panels (Schiffbaenker et al., 2022; van den Besselaar et al., 2018). For this reason, it is highly necessary to explore alternative explanations for this finding, for which we propose three: bias awareness leading to overcompensation, perceived scrutiny, and stereotype reversal. Our interpretation makes the assumption that participants were aware of the existence of such a bias, which is not unreasonable given that this is an amply discussed topic and is known to most members of academia. Based on this awareness, any of these three effects or a combination of them could have occurred. Regarding overcompensation, panelists, becoming aware of their bias toward male candidates, might have attempted to consciously correct this but overcompensated, leading to a new bias toward female candidates. This phenomenon has been known to occur in other contexts (e.g., Chung et al., 2001; Madden et al., 2021), so it would not be surprising to find it

in grant allocation as well. Regarding perceived scrutiny, as participants knew that the results of the study were intended for publication, it stands to reason that social desirability (Fisher, 1993) influenced their responses. Regarding stereotype reversal, there could be a perceived need to “balance the scales,” leading to an overemphasis on selecting female candidates (Bol et al., 2022). Some participants offered feedback on the survey debriefing that is aligned with these interpretations. Although we will not transcribe their comments here for privacy, a few participants explicitly stated using gender as a criterion to correct the existing gender gap, while one participant mentioned needing to obfuscate the part of the screen where gender and age were presented to avoid triggering a biased response. Thus, regardless of the actual direction of the bias, this finding suggests that remedial action is indeed necessary in panel evaluations.

The impact of age is complex since it can certainly be due to ageism, a bias toward age (Palmore, 1999). However, a more benevolent interpretation is that panelists prefer to extend an opportunity to younger researchers who are starting their careers rather than allocating funds to researchers who are much later in their careers. This can be corroborated by the findings in the narrative persona analysis, in which the Underdog, a young underachiever with a great project, was chosen much more frequently than the Fundraiser, an older researcher who had already secured a number of grants. There is limited literature on age biases in academia, but academic productivity tends to decrease over the career life cycle as tenure is secured (and as such papers lose some of the value that they have for career progression) and priorities shift to other obligations (Dunn, 2005). Thus, the negative impact of age could also be an indication that the participants were aware of this fact and responded accordingly by favoring younger researchers, who are more likely to benefit from the funding.

A serendipitous and interesting finding was the fact that, despite multiple mechanisms in place to ensure the numerical balance of the two personas, the derivative persona was chosen substantially more often than the seed persona. As diagnosis modeling revealed that this was not due to the larger range of values, the remaining explanation is that this was caused by the higher dispersion of values in the derivative persona. Nevertheless, it is remarkable that this had an effect. As the order of presentation was shuffled and participants saw only 10 personas in total (in five pairs), it is very unlikely that participants were able to consciously deduce the higher dispersion of values based on only the provided information. The author and colleagues were unable to identify which persona was which in a blind trial, even when they were aware of the generation mechanism. Regardless, the only remaining plausible explanation given the available data is that panelists prefer candidates with less dispersed attributes, which is corroborated by the effect of our measure of persona variability – personas which were generated with a smaller dispersion of attributes were more likely to be chosen. One possible interpretation is that panelists prefer candidates which have a balance of attributes in their profiles, rather than candidates which exhibit more extreme fluctuations or have uneven profiles, such as having excellent bibliometric indicators but lacking other types of experience (or vice-versa). In this sense, this could be an indication that attempting to gamify evaluation systems by focusing on inflating just a few criteria to the detriment of others can be counter-productive.

Finally, we will not delve much into the pairwise trait comparisons for the narrative persona analysis, as they were already discussed throughout their presentation. However, we will comment on the overall picture of the optimal candidate. It portrays a relatively prolific mid-career researcher with a balance of traits. There is a clear preference for a multidisciplinary approach, which is sensible given the current research panorama (Martimianakis & Muzzin, 2015), and this goes hand in hand with a preference for collaborative endeavors, with just the right amount of solo work, likely to

demonstrate the capability for autonomy. The optimal candidate is also focused on applied science, which is not surprising given that this is considered highly preferable in many evaluation schemes, notably when it comes to outputs with commercial applications (de Jong et al., 2016; Smit & Hessels, 2021). This optimal candidate is also capable of balancing, in an equal manner, interactions with industry players and the general public; this is a rare combination of traits but one that is certainly possible and quite remarkable when it presents itself. The importance of the project, which was presented as a feature in only one of the pairwise choices, is also not to be understated, as our fictional Underdog was able to beat a more seasoned researcher while having only an excellent project to his name. Finally, our optimal candidate should be outspoken and unafraid to challenge established norms, a trait inherited from the Iconoclast in our pairwise comparison. This suggests at least some degree of discontent with currently established scientific practices. Although we deliberately left some room for interpretation in this comparison, feedback from the participants indicated some discontent with grant evaluation practices, which is unsurprising given the recent emergence of underlying issues with most systems (e.g., Jerrim & Vries, 2020; Roumbanis, 2021) and a shift, for example, to lottery-based funding (Smaldino et al., 2019).

4.1. Limitations

This exercise aimed to study a process for which data is scarce, making use of fictional personas to surmount this obstacle. However, although the quantitative personas were driven by real-world data and the narrative personas were grounded in existing literature, these are still simplified versions of what a real-world candidate would look like. For this reason, this study should be taken for what it is - a representation of the general preferences of panelists given the available information, and not a complete modeling of panelist choice behavior, which is unlikely to ever be developed, given that most of the data required for such an exercise is confidential. With that said, an argument that could be used against the present study - that many of the choices had to be made under uncertainty, as only limited information on the fictional candidates was made available - can be countered by the fact that in real-world scenarios, such choices are, arguably, also made under uncertainty, as panelists seldom have all possible information that could potentially matter for such decisions (Zhu et al., 2022).

This leads to a second important limitation, which is the focus on the candidate and the absence of a fictional project to simulate the entire contents of an actual application. This approach, albeit ideal, would not be realistically feasible: such a survey would be too time-consuming for most potential participants; tailor-made fictional projects would need to be created for each field of science, potentially at a topical level, so that participants could make meaningful evaluations; and the analysis itself would be extremely difficult to implement due to the inherent subjectivity in evaluating a project. Consequently, a deliberate decision was made to discard the project component and present the exercise to the participants as a series of tiebreaks between candidates with similar scores on the project component. Additionally, the projects could be either basic or applied, so it would be plausible that all candidates had an alignment with the proposal. An argument can be made that the project is the only component that really matters in grant applications. To this, we present two counterarguments: (a) in some grant programs (e.g., ERC), the candidate's profile is an explicit evaluation criteria beyond the project, and (b) the literature has already established the importance of the candidate's profile on selection processes even when it is not part of the explicit criteria (Lamont, 2009). Thus, despite the absence of a project being a limitation, we nevertheless argue that studying the candidate's features is still relevant, and we posit that they continue to have an impact in the presence of an accompanying project. The question that

remains, then - and one that, unfortunately, cannot be answered with our approach - is how much of an impact they have when presented together with such a project.

4.2. Implications

This study has several important implications. The unexpectedly large effect of gender and age, when compared to the other career-based metrics, suggests that the panelists made a conscious effort to avoid bias and, in doing so, swung too far in the other direction. One way to address this issue is to provide training to panelists on unconscious bias, emphasizing that the goal is not to favor one group over another but to evaluate each candidate on their merits. A focus on increasing awareness of the existence and impact of unconscious biases, as well as providing strategies to minimize their influence, might help ensure a fairer and more equitable selection process. It is also important to note the importance of continuous evaluation and adjustment in attempts to rectify biases. Changes in behavior following interventions should be monitored not only for the elimination of the original bias but also to determine whether overcorrection or the emergence of new biases has occurred. Bias correction is a nuanced process, and the aim should be to achieve fairness and objectivity, not merely to reverse the direction of bias. An alternative approach would be to simply fully blind applications so that irrelevant candidate information is not accessible. The problem with this approach is that, due to the nature of publication track records, a panelist who is so inclined can still perform an online search - for example, for an article in the track record - and easily find the candidate's information. However, this would still be preferable to immediately providing unnecessary information, and it is reasonable to assume that panelists, given their immense workload, are unlikely to perform individual background checks on every candidate. Finally, a third option would be to adopt lottery-based systems, which were already mentioned in this article, and are a more practical approach of solving this issue – and one which is already being used in some schemes.

This study also has important implications for academics applying to grant programs by providing information on what an optimal candidate would look like. Although this information is likely to be of little use to seasoned researchers who have already shaped and established their careers, it can be useful for new doctorates, as it provides hints regarding the path their careers should take if their aim is to secure research funding as soon as possible. Notably, a recommendation would be to focus on publications but also to foster both industry and public outreach while establishing multidisciplinary collaborative networks.

A final implication is on a methodological level. To the best of our knowledge, this is the first study to apply the use of personas to the study of panel evaluation processes. In this sense, it can also be seen as a proof of concept for how this approach can be used to tackle topics for which there is little available information. In particular, the procedurally generative approach that was employed, coupled with multi-level modeling, is particularly powerful, since it can be used to generate a very large data set with relatively few participants. As the methodology further develops, more features can be added, which are certain to produce a wealth of interesting studies in the near future.

5. References

- Abramo, G., D'Angelo, C. A., & Rosati, F. (2016). Gender bias in academic recruitment. *Scientometrics*, 106(1), 119–141. <https://doi.org/10.1007/s11192-015-1783-3>
- Alberts, B., Hanson, B., & Kelner, K. L. (2008). Reviewing peer review. *Science*, 321(5885), 15–15.

- Angervall, P. (2018). The academic career: A study of subjectivity, gender and movement among women university lecturers. *Gender and Education*, 30(1), 105–118.
<https://doi.org/10.1080/09540253.2016.1184234>
- Arora, A., David, P. A., & Gambardella, A. (1998). Reputation and Competence in Publicly Funded Science: Estimating the Effects on Research Group Productivity. *Annales d'Économie et de Statistique*, 49/50, 163–198. <https://doi.org/10.2307/20076114>
- Ayoubi, C., Pezzoni, M., & Visentin, F. (2021). Does It Pay to Do Novel Science? The Selectivity Patterns in Science Funding. *Science and Public Policy*, 48(5), 635–648.
<https://doi.org/10.1093/scipol/scab031>
- Banal-Estañol, A., Macho-Stadler, I., & Pérez-Castrillo, D. (2019). Evaluation in research funding agencies: Are structurally diverse teams biased against? *Research Policy*, 48(7), 1823–1840.
<https://doi.org/10.1016/j.respol.2019.04.008>
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*. McGraw-Hill International.
- Bol, T., de Vaan, M., & van de Rijt, A. (2022). Gender-equal funding rates conceal unequal evaluations. *Research Policy*, 51(1), 104399. <https://doi.org/10.1016/j.respol.2021.104399>
- Chung, Y. B., Marshall, J. A., & Gordon, L. L. (2001). Racial and Gender Biases in Supervisory Evaluation and Feedback. *The Clinical Supervisor*, 20(1), 99–111.
https://doi.org/10.1300/J001v20n01_08
- Conceição, P., Heitor, M. V., Sirilli, G., & Wilson, R. (2004). The “swing of the pendulum” from public to market support for science and technology: Is the U.S. leading the way? *Technological Forecasting and Social Change*, 71(6), 553–578.
<https://doi.org/10.1016/j.techfore.2004.03.002>
- de Jong, S. P. L., Smit, J., & van Drooge, L. (2016). Scientists’ response to societal impact policies: A policy paradox. *Science and Public Policy*, 43(1), 102–114.
<https://doi.org/10.1093/scipol/scv023>

- Dunn, J., Robert. (2005). The age bias in academic publishing. *Challenge*, 48(5), 5–11.
- Fisher, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research*, 20(2), 303–315. <https://doi.org/10.1086/209351>
- Hamlyn, B., Shanahan, M., Lewis, H., O'Donoghue, E., Hanson, T., & Burchell, K. (2015). *Factors affecting public engagement by researchers*.
- Haslam, N., & Laham, S. (2009). Early-career scientific achievement and patterns of authorship: The mixed blessings of publication leadership and collaboration. *Research Evaluation*, 18(5), 405–410. <https://doi.org/10.3152/095820209X481075>
- Hilmer, C. E., & Hilmer, M. J. (2005). How Do Journal Quality, Co-Authorship, and Author Order Affect Agricultural Economists' Salaries? *American Journal of Agricultural Economics*, 87(2), 509–523. <https://doi.org/10.1111/j.1467-8276.2005.00738.x>
- Jansen, B. J., Salminen, J., Jung, S., & Guan, K. (2022). *Data-driven personas*. Springer Nature.
- Jerrim, J., & Vries, R. de. (2020). Are peer-reviews of grant proposals reliable? An analysis of Economic and Social Research Council (ESRC) funding applications. *The Social Science Journal*, 1–19.
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: An introduction to cluster analysis* (Vol. 344). John Wiley & Sons.
- Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago press.
- Lamont, M. (2009). *How Professors Think: Inside the Curious World of Academic Judgment*. Harvard University Press.
- Larivière, V., Macaluso, B., Archambault, É., & Gingras, Y. (2010). Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation*, 19(1), 45–53.
- Laudel, G. (2006). The 'Quality Myth': Promoting and Hindering Conditions for Acquiring Research Funds. *Higher Education*, 52(3), 375–403. <https://doi.org/10.1007/s10734-004-6414-5>

- Lauto, G., & Valentin, F. (2013). How Large-Scale Research Facilities Connect to Global Research. *Review of Policy Research*, 30(4), 381–408. <https://doi.org/10.1111/ropr.12027>
- Lawson, C., & Salter, A. (2023). Exploring the effect of overlapping institutional applications on panel decision-making. *Research Policy*, 52(9), 104868. <https://doi.org/10.1016/j.respol.2023.104868>
- Ma, L., & Agnew, R. (2022). Deconstructing impact: A framework for impact evaluation in grant applications. *Science and Public Policy*, 49(2), 289–301. <https://doi.org/10.1093/scipol/scab080>
- Madden, D., Liu, Y., Yu, H., Sonbudak, M. F., Troiano, G. M., & Hartevelt, C. (2021). “Why Are You Playing Games? You Are a Girl!”: Exploring Gender Biases in Esports. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. <https://doi.org/10.1145/3411764.3445248>
- Marsh, H. W., & Hattie, J. (2002). The Relation between Research Productivity and Teaching Effectiveness: Complementary, Antagonistic, or Independent Constructs? *The Journal of Higher Education*, 73(5), 603–641.
- Martimianakis, M. A., & Muzzin, L. (2015). Discourses of interdisciplinarity and the shifting topography of academic work: Generational perspectives on facilitating and resisting neoliberalism. *Studies in Higher Education*, 40(8), 1454–1470.
- Martin, B. R. (2011). The Research Excellence Framework and the ‘impact agenda’: Are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247–254.
- Mazzucato, M. (2013). *The entrepreneurial state: Debunking public vs. Private sector myths* (Vol. 1). Anthem Press.
- Millar, M. M. (2013). Interdisciplinary research and the early career: The effect of interdisciplinary dissertation research on career placement and publication productivity of doctoral graduates in the sciences. *Research Policy*, 42(5), 1152–1164. <https://doi.org/10.1016/j.respol.2013.02.004>

- Mongeon, P., Brodeur, C., Beaudry, C., & Larivière, V. (2016). Concentration of research funding leads to decreasing marginal returns. *Research Evaluation*, 25(4), 396–404.
- Murray, D. L., Morris, D., Lavoie, C., Leavitt, P. R., MacIsaac, H., Masson, M. E. J., & Villard, M.-A. (2016). Bias in Research Grant Evaluation Has Dire Consequences for Small Universities. *PLOS ONE*, 11(6), e0155876. <https://doi.org/10.1371/journal.pone.0155876>
- Öcalan-Özel, S., & Llerena, P. (2021). Industry Collaborations of Research Teams: Are They Penalized or Rewarded in the Grant Evaluation Process? *Frontiers in Research Metrics and Analytics*, 6. <https://www.frontiersin.org/articles/10.3389/frma.2021.707278>
- Palmore, E. (1999). *Ageism: Negative and Positive, 2nd Edition*. Springer Publishing Company.
- Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., Ford, C. E., & Carnes, M. (2018). Low agreement among reviewers evaluating the same NIH grant applications. *Proceedings of the National Academy of Sciences*, 115(12), 2952–2957.
- Rao, V. R. (2014). *Applied Conjoint Analysis*. Springer. <https://doi.org/10.1007/978-3-540-87753-0>
- Reed, M. S., Duncan, S., Manners, P., Pound, D., Armitage, L., Frewer, L., Thorley, C., & Frost, B. (2018). A common standard for the evaluation of public engagement with research. *Research for All*, 2, 143–162. <https://doi.org/10.18546/RFA.02.1.13>
- Roumbanis, L. (2021). Disagreement and Agonistic Chance in Peer Review. *Science, Technology, & Human Values*, 01622439211026016. <https://doi.org/10.1177/01622439211026016>
- Rowe, G., & Frewer, L. J. (2004). Evaluating public-participation exercises: A research agenda. *Science, Technology, & Human Values*, 29(4), 512–556.
- Salminen, J., Liu, Y.-H., Şengün, S., Santos, J. M., Jung, S., & Jansen, B. J. (2020). The effect of numerical and textual information on visual engagement and perceptions of AI-driven persona interfaces. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 357–368. <https://doi.org/10.1145/3377325.3377492>

- Santos, J. M. (2022). Quis judicabit ipsos judices? A case study on the dynamics of competitive funding panel evaluations. *Research Evaluation*, rvac021.
<https://doi.org/10.1093/reseval/rvac021>
- Santos, J. M., Horta, H., & Amâncio, L. (2020). Research agendas of female and male academics: A new perspective on gender disparities in academia. *Gender and Education*, 1–19.
<https://doi.org/10.1080/09540253.2020.1792844>
- Santos, J. M., Horta, H., & Li, H. (2022). Are the strategic research agendas of researchers in the social sciences determinants of research productivity? *Scientometrics*.
<https://doi.org/10.1007/s11192-022-04324-7>
- Schiffbaenker, H., van den Besselaar, P., Holzinger, F., Mom, C., & Vinkenburg, C. (2022). Gender Bias in Peer Review panels. *Inequalities and the Paradigm of Excellence in Academia*, 109.
- Seeber, M., Alon, I., Pina, D. G., Piro, F. N., & Seeber, M. (2022). Predictors of applying for and winning an ERC Proof-of-Concept grant: An automated machine learning model. *Technological Forecasting and Social Change*, 184, 122009.
<https://doi.org/10.1016/j.techfore.2022.122009>
- Smaldino, P. E., Turner, M. A., & Contreras Kallens, P. A. (2019). Open science and modified funding lotteries can impede the natural selection of bad science. *Royal Society Open Science*, 6(7), 190194. <https://doi.org/10.1098/rsos.190194>
- Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: A review of societal impact assessment methods. *Research Evaluation*, 30(3), 323–335. <https://doi.org/10.1093/reseval/rvab002>
- Tight, M. (2016). Examining the research/teaching nexus. *European Journal of Higher Education*, 6(4), 293–311. <https://doi.org/10.1080/21568235.2016.1224674>
- van den Besselaar, P., Schiffbaenker, H., Sandström, U., & Mom, C. (2018). *Explaining gender bias in ERC grant selection—Life Sciences case*. 346–352.

Weenink, K., Aarts, N., & Jacobs, S. (2023). 'I need a grant but spend time on teaching': How academics in different positions play out the teaching–research nexus in interdependence with their contexts. *European Journal of Higher Education*, 0(0), 1–19.

<https://doi.org/10.1080/21568235.2023.2221839>

Wray, K. B. (2002). The Epistemic Significance of Collaborative Research. *Philosophy of Science*, 69(1), 150–168. <https://doi.org/10.1086/338946>

Zhang, T., Ramakrishnon, R., & Livny, M. (1996). BIRCH: An efficient data clustering method for very large data bases. 1996. *Montreal, Canada*, 103–114.

Zhu, W., Li, S., Zhang, H., Zhang, T., & Li, Z. (2022). Evaluation of scientific research projects on the basis of evidential reasoning approach under the perspective of expert reliability. *Scientometrics*, 127(1), 275–298. <https://doi.org/10.1007/s11192-021-04201-9>

List of Figures

Figure 1. Distribution of probabilities by cluster class.

Figure 2. Persona preferences for each pair.

Figure 3. Global persona preferences.