Working Paper 23-037

The Subjective Expected Utility Approach and a Framework for Defining Project Risk in Terms of Novelty *and* Feasibility—A Response to Franzoni and Stephan (2023), 'Uncertainty and Risk-Taking in Science'

Jacqueline N. Lane

The Subjective Expected Utility Approach and a Framework for Defining Project Risk in Terms of Novelty *and* Feasibility—A Response to Franzoni and Stephan (2023), 'Uncertainty and Risk-Taking in Science'

Jacqueline N. Lane
Harvard Business School

Working Paper 23-037

Funding for this research was provided in part by Harvard Business School.

The Subjective Expected Utility Approach and a Framework for Defining Project Risk in Terms of Novelty *and* Feasibility – A Response to Franzoni and Stephan (2023), 'Uncertainty and Risk-Taking in Science'

Jacqueline N. Lane

Harvard Business School
Harvard University
Morgan Hall 428
Boston, MA 02163, USA
jnlane@hbs.edu

January 3, 2023

Abstract

In their Discussion Paper, Franzoni and Stephan (F&S, 2023) discuss the shortcomings of existing peer review models in shaping the funding of risky science. Their discussion offers a conceptual framework for incorporating risk into peer review models of research proposals by leveraging the Subjective Expected Utility (SEU) approach to decouple reviewers' assessments of a project's potential value from its risk. In my Response, I build on F&S's discussion and attempt to shed light on three additional yet core considerations of risk in science: 1) how risk and reward in science are related to assessments of a project's novelty and feasibility; 2) how the sunk cost literature can help articulate why reviewers tend to perceive new research areas as riskier than continued investigation of existing lines of research; and 3) how drawing on different types of expert reviewers (i.e., based on domain and technical expertise) can result in alternative evaluation assessments to better inform resource allocation decisions. The spirit of my Response is to sharpen our understanding of risk in science and to offer insights on how future theoretical and empirical work—leveraging experiments— can test and validate the SEU approach for the purposes of funding more risky science that advances the knowledge frontier.

This version has been accepted for publication in Research Policy but has not been copyedited by the journal.

1. Introduction

In recent years, innovation scholars, program officers and policymakers alike, have grown increasingly concerned with conservatism in funding decisions in science—a pattern in which the evaluation outcomes of scientific peer review processes tend to be biased against the funding of risky research projects. A direct implication of conservatism is that it can slow progress on pioneering and innovative research studies essential to scientific and technological breakthroughs. In this discussion paper, Franzoni and Stephan (F&S, 2023) seek to address this critical concern by proposing an alternative review model—the Subjective Expected Utility (SEU) framework—that offers a systematic approach for incorporating risk into the peer review process by requiring reviewers to make separate assessments of a project's potential risk and value. F&S motivates the SEU framework by first offering a detailed discussion of the meaning of risk in science and adjacent literature, and then considering three sources of uncertainty that are likely to affect funding decisions within science: (i) uncertainty in outcomes, (ii) uncertainty in the probability of discovery, and (iii) uncertainty in the value of findings. By highlighting the central role that risk plays in the peer review process for grant funding decisions, F&S offers a constructive lens for increasing our scholarly understanding of how and when risk-taking in science can advance the knowledge frontier.

In my Response to F&S, I focus my remarks primarily on the SEU as a pragmatic yet promising alternative peer review model for funding risky science. My commentary is structured as follows: first, I will begin by commenting on the strengths of the SEU approach over existing models for evaluating scientific research proposals; second, I will turn to three components of F&S's discussion that warrant greater attention and refinement to sharpen the SEU approach. These three components are (i) characterizing scientific risk and reward in terms of novelty *and* feasibility, (ii) incorporating sunk cost theory into perceptions of project risk and/or probability of success, and (iii) leveraging alternative forms of expertise (e.g., domain and technical) to better inform the inputs in the SEU.

The SEU approach to funding research proposals has given me much food for thought. One key reason to advocate for the SEU framework over traditional peer review models is that it marks a notable shift away from basing funding decisions on a single overall score of quality. In traditional peer review models, such as those used by the U.S. National Institutes of Health (NIH) or National Science Foundation (NSF), reviewers need to make interpretative decisions about how to weight the relative importance of different peer review criteria in their assessments of a proposal's overall quality. Because there is no single, objective weight function that determines how these criteria should be jointly applied, the aggregation process can introduce bias into the peer review process. Reviewers may prioritize different criteria and use inconsistent weighting schemes to arrive at their evaluation scores (Lee, 2015), or they may be negatively influenced by others' opinions (Lane et al., 2021b). This can be problematic because it can result in funding decisions that depend more on the reviewer to whom the grant application was assigned rather than the proposed research (Pier et al., 2018). In addition, although most reviewers tend to incorporate a project's risk into their evaluation scores, standard peer review models do not ask directly about a project's riskiness—suggesting that reviewers may not even be aware that risk is being implicitly factored into their weighting of component criteria (Gallo et al., 2018).

Instead, F&S's proposed SEU framework disaggregates the overall quality score into distinct components of value and risk, with the combined outcome being computed by multiplying the parts together. Reviewers need to consider the primary and secondary outcomes of a potential project, as well as the value and probabilities associated with each outcome. This approach is promising for encouraging risky science because it requires reviewers to make direct assessments of a project's risk—a departure from current peer review approaches—which I agree is critical for overcoming conservatism in funding decisions. I believe the SEU approach provides the building blocks for a promising peer review model for assessing research proposals going forward. Before we get there, we need to do more to clarify, test, and validate the SEU. Below I outline three areas that warrant greater investigation to further develop and flesh out the SEU approach. I then suggest that field experiments represent a compelling way to test hypotheses about the applicability of the SEU for funding risky science.

2. Relating Risk and Reward to Project Novelty and Feasibility

As F&S's discussion of risk reveals, the meaning of risk in science has several dimensions, which include uncertainty with respect to outcomes, probabilities, and values. Yet despite this classification, the concept of risk may still be too abstract to be implemented into the SEU for use as a peer review model. If we examine the sub-questions in Table 1 (F&S, 2023, p. 20), the authors propose several relevant questions for investigating the underlying risk of a proposal. However, it is unclear how these questions might be translated into values and probabilities to inform the SEU. In the current SEU, reviewers are required to come up with the value of each outcome as well as the subjective probability associated with each outcome, which is computed as the product of their natural and methodological probabilities. I see this as problematic due to two underlying assumptions about a project's risk-reward profile that were revealed through F&S's discussion of the prior literature on science funding decisions. The first assumption is that risk is often conflated with novelty even though they are different constructs. The second assumption is that the relationship between risk and reward, as well as that between novelty and feasibility, is often perceived in terms of tradeoffs in the proposed research, even though there may be other kinds of relationships between these two project dimensions.

In this section, I discuss why these assumptions are problematic for funding risky science. I then suggest an alternative conceptualization for investigating the relationship between the value and probability of success of research proposals that disaggregates a project's risk by its novelty and feasibility. In this alternative conceptualization, reviewers would first make separate assessments about a project's novelty and feasibility, which would then be used to inform their judgments about a project's risk and reward characteristics. As I discuss below, this alternative conceptualization would allow reviewers to disentangle novelty from risk, as well as perceive a range of possible relationships between risk and reward in a project's characteristics based on their assessments of its novelty and feasibility.

First, project novelty and risk are often conflated in science even though they are distinct constructs. As Kuhn (1977) recognized, there is an essential tension between rewarding novelty and upholding the body of established knowledge. This is echoed by Hackett and Chubin (2003) who suggest that there is a difference between "sound innovation and reckless speculation" (ibid., p. 10). Hence, a primary objective of the peer review process is to systematically differentiate novel ideas (i.e., sound innovation) from bizarre or fanciful ones (i.e., reckless speculation). For these reasons, it is critical that the SEU appropriately differentiates novel research from risk to come up with unbiased assessments of the research's value. While novel research has the potential to yield high value, reckless speculation, or risk, would not. An important consideration, however, is whether the current SEU framework is able to disentangle novelty from risk—a task whose difficulty may be further amplified given that novelty as a concept can be abstract, context dependent, and in the eye of the beholder (Mount et al., 2021).

Second, risk and reward are often considered to be tradeoffs in science, meaning that risky research is assumed to yield high reward. To illustrate this point, in their discussion of the prior literature, F&S indicate that funding agencies tend to focus "excessively on feasibility and things that can go wrong at the expense of supporting research that has the potential to lead to major breakthroughs" (F&S, 2023, p. 3), or that reviewers are prone to "under-reacting to potential gains, and over-reacting to potential problems, or anchoring evaluations to easy reference points" (ibid., p. 24). These statements highlight that this tradeoff view in a project's risk-reward profile is related to reviewers' perceptions of novelty and feasibility as also being tradeoffs in a project's characteristics. In other words, a takeaway from this discussion is that we would expect highly novel work to be lower in feasibility—i.e., high risk and high reward, and highly feasible work to be lower in novelty—i.e., low risk and low reward.

I contend that this tradeoff view is only one possible relationship between a project's novelty and feasibility, as well as its risk-reward profile. To better inform the inputs of the SEU, we need a more systematic understanding of how a project's risk-reward profile is related to its novelty and feasibility. In turn, I ask: *Under what conditions might the novelty and feasibility of a proposed research project be tradeoffs and when might the relationship between these dimensions be more one of complementarity?*

In Figure 1, I propose a framework for characterizing risk and reward in science grant funding in terms of the relationships between the novelty and feasibility of the research project. The x-axis is a project's feasibility, which describes how hard or easy a project is to implement. The y-axis is a project's novelty, which describes its originality or the degree that it departs from existing knowledge in the research area or field.[1] Although F&S discuss novelty and feasibility in the sub-questions listed in Table 1 (p. 20), Figure 1 investigates how different novelty-feasibility configurations, as perceived by the reviewer, might be incorporated into a framework for describing a project's risk and reward characteristics. Hence, Figure 1 provides greater specificity into F&S's observation that risk and reward need not be correlated in science (p. 6). Framed in this way, *Grand Challenges* and *Low-hanging-fruit* correspond to extreme points, where project novelty moves from high to low and project feasibility simultaneously moves from low to high.[2] *Grand Challenges*, defined as difficult yet pivotal problems—often associated with "breakthrough" innovations—tend to be high risk, high reward projects. Examples of current *Grand Challenges* include the NIH, DARPA, and NSF's BRAIN initiative to uncover new ways to treat, prevent and cure brain disorders, such as Alzheimer's, autism, and traumatic brain injury, and DOE's SunShot *Grand Challenge* to make solar energy's cost competitive with coal within the decade.[3] The *Low-hanging-fruit* correspond to easy-to-solve, straightforward problems. These projects might include extending research lines on existing datasets, improving experimental designs, or conducting replications of past experiments to improve their generalizability. In contrast to *Grand Challenges*, *Low-hanging-fruit* tend to be low risk yet low reward projects that can be carried out with relatively little uncertainty.

The other endpoints of the framework correspond to *Difficult* problems and *Pivotal* problems for which the dimensions of novelty and feasibility are complementary characteristics. *Difficult problems* correspond to projects that are low novelty and low feasibility. At first glance,

[1] See Stokes' *Pasteur's Quadrant* for an alternative two-dimensional classification of scientific research projects, with one axis being "inspired by a quest for fundamental understanding" and the other being "inspired by considerations of use" (Stokes, 1997)
[2] See Alon (2009) who uses the same terminology to define research projects at similar extreme end points.
[3] https://obamawhitehouse.archives.gov/administration/eop/ostp/grand-challenges

it may not be entirely obvious why researchers would pursue *Difficult* problems, but history indicates that such problems exist. One example of a notoriously difficult problem was Fermat's Last Theorem, which took 350 years to solve. Fermat claimed that $x^n + y^n = z^n$ has no non-zero integer solutions for x, y and z when $n > 2$. The problem had been unsolved for 300 years—and was no longer considered novel to mathematicians—by the time the young mathematician, Andrew Wiles came across it. In fact, solving it was no longer viewed as being exceptionally important to the field of mathematics at the time (Lew, 2005). After studying elliptical curves at the University of Cambridge, Wiles worked on the problem in secrecy and isolation for seven years before he finally finished the proof in 1994. In recognition of this achievement, Wiles was later awarded the 2016 Abel Prize and 2017 Copley Medal by the Royal Society.[4]

Fermat's Last Theorem illustrates that a project's risk can be driven by its (low) feasibility rather than its novelty. Fermat's Last Theorem had stumped many mathematicians and was assumed to be intractable, but by no means a novel problem by the time Wiles applied elliptical curves theory to solve it. Therefore, Fermat's Last Theorem and similar problems are high risk—due to their lack of proven approach or methodology—but with unknown-reward because their *ex ante* value to science and/or society is uncertain.

The last endpoint corresponds to *Pivotal problems*, which are characterized as both high novelty and high feasibility. *Pivotal* problems are projects for which the researcher's choice of *perspective* contributes to a problem's difficulty (Page, 2008). Although a proposed project may appear difficult or unsolvable to some individuals, it may be solvable with relative ease by others who bring an alternative perspective from another field or domain of expertise (Jeppesen and Lakhani, 2010; McLaughlin, 2001) that ultimately makes the problem feasible to be solved. For

[4] https://www.britannica.com/biography/Andrew-Wiles

example, many physicists—including Francis Crick and Sir John Randall—seized the opportunity to enter the field of biology at the end of World War II, because they observed that their expertise in physics could be applied in meaningful ways to solve *Pivotal* problems at the forefront in biology. The impact on biology from these physicists has been immeasurable, with many of them winning Nobel prizes for their achievements.[5] Funding agencies, such as the NIH, have also recognized the benefit of enticing "outsiders" or "new brains" with alternative training, backgrounds, and expertise into a field to solve pivotal problems called Requests for Applications (RFAs) using large monetary grants of $2 to $3 million (Myers, 2020). In short, *Pivotal problems* tend to have unknown risk—because their ability to be solved depends largely on identifying individuals that bring a potentially fruitful perspective—but are likely to bring high reward if solved.

This characterization of project risk in terms of novelty and feasibility presented in Figure 1 has the potential to sharpen our understanding of the SEU framework because it decouples risk from reward, by instead bringing to the forefront considerations of a project's novelty and feasibility. Another insight of Figure 1 is that it suggests that a project's risk can be driven by its novelty, feasibility or by both its novelty *and* feasibility. Whereas the risk associated with *Difficult* problems is primarily driven by coming up with a feasible approach, the risk associated with *Pivotal* problems lies primarily in its novelty relative to the knowledge frontier. Lastly, *Grand Challenges* are risky because they have characteristics of both pivotal and difficult problems, making them both highly novel and difficult to solve due to their low feasibility. While I focus on four extreme points to illustrate the potential relationships between novelty and feasibility, it is important to note that most projects will fall somewhere within these extremes. The goal of the

[5] https://www.nobelprize.org/prizes/medicine/1962/crick/biographical/

review process would be to recruit reviewers to evaluate where a project's novelty and feasibility lie with respect to the framework depicted in Figure 1, and then make assessments of its risk-reward profile, accordingly.

Going further and beyond the SEU, I assert that our understanding of funding risky science can be improved by paying greater attention to a project's feasibility, in addition to its novelty. Most literature to date has tended to focus on the dimension of novelty as a project characteristic but has paid relatively less attention to the feasibility dimension. Growing evidence suggests that science may be "anti-novelty", as highly novel ideas either face resistance from incumbents in the field or are delayed in terms of recognizing their contribution to science and/or society compared to less novel ideas (Boudreau et al., 2016; Wang et al., 2017). Yet it is also possible that conservatism in science may under certain conditions be driven by a *feasibility preference* in lieu of an *anti-novelty preference*. Disentangling these two types of preferences from one another has critical implications for how project risk is articulated and how resources might be allocated to projects. For example, whereas a tendency to select feasible projects for funding might signal a preference for existing approaches and methodologies, a tendency to select less-novel projects might signal a preference for research that is grounded in the existing literature and is consistent with the core views in a field or domain. Such documented preferences for less-novel or incremental projects is consistent with the notion that novelty tends to be associated with greater uncertainty in reviewer judgments, which may lead to lower scores for novel research (Luukkonen, 2012; Mueller et al., 2012). Going forward, a fruitful direction may be to examine how different types of novelty proposed in a research project (e.g., theory, application context, method, statistical analyses) impact assessments of risk and reward—and the degree to which certain types of novelty may be better suited for the SEU framework.

Shedding light on reviewers' preferences for different types of novel and/or feasible work echoes the comments raised in the other Reponses by Heinze (this issue), who makes the point, using Whitley (2000)'s typology of technical and strategic task uncertainties that certain disciplines, such as physics or chemistry may be better suited for the SEU framework due to the predictability of task outcomes and problems, as well as by Stirling (this issue) who argues that the risks associated with some research activities may not be measurable. Against this backdrop, the innovation literature has revealed that novelty tends to be an abstract concept; its ability to be recognized is likely to depend on factors, such as the reviewers' depth of knowledge in the domain (Boudreau et al., 2016; Mount et al., 2021) or their position as a core or peripheral member (Azoulay et al., 2019; McLaughlin, 2001). Although the framework depicted in Figure 1 portrays novel research based on the originality of the contribution to the field or discipline, it is important to note that reviewers tend to hold different thresholds for recognizing novelty (e.g., new-to-the-subfield vs. new-to-the-world) (Azoulay et al., 2019; Uzzi et al., 2013). Therefore, assessments of risk and reward may also be colored by one's perceptions of novelty. Accordingly, the SEU framework may be more likely to apply in fields where there is greater consensus over perceptions of novelty or alternatively, reserved for research problems that yield consensus about their novelty but are driven by the difficulty along the feasibility dimension (i.e., *Difficult Problems*). Because feasibility tends to be a more concrete characteristic of a research problem, given a set of parameters or specifications, it may lend itself to be more naturally measured.

Ultimately, a deeper understanding of risk and reward—in terms of novelty and feasibility—is useful for project funding and selection decisions as it will not only improve the SEU framework but will also be valuable for designing approaches to train and inform reviewers so that they are more attuned to the resource allocation goals of the funding opportunity they are evaluating and of their own cognitive limits.

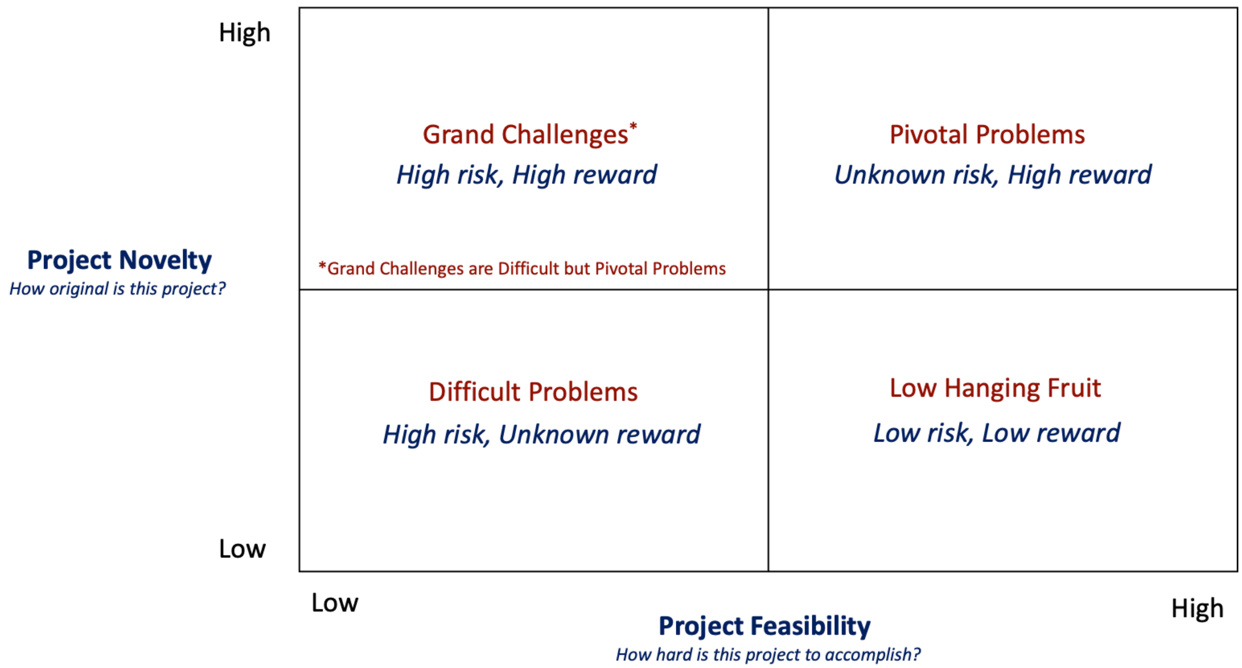

Figure 1. A framework for defining research projects.

3. Incorporating Sunk Costs into the SEU

In F&S's discussion of risk-taking in science, I found their anecdote of James P. Allison's discovery of immune checkpoint inhibitors in cancer treatment to be particularly thought-provoking. F&S use this anecdote to illustrate a common misperception within science, where it is often assumed that engaging in a new line of research is the riskier course of action compared to continuing along an existing research path (F&S, 2023, p. 8). As F&S write: "persisting along a line of research can often provide a more predictable path, albeit one with diminishing returns, or one potentially leading to a dead-end" (ibid., p. 9). This observation regarding the comparative risk versus reward of taking on a new research area versus continuing down an existing research line—albeit with little upside—warrants greater attention in the SEU approach. In this section, I unpack this tension between the relative attractiveness of funding new versus existing research areas and discuss the role of sunk costs in potentially shaping this common misperception to invest in existing research even when a new research line may be the more attractive option. I suggest that sunk costs, in terms of prior investments in an existing research path—may inflate a reviewer's perception of a project's probability of success or p(s). In this section, I discuss how and why inflation of p(s) may occur and how it might affect assessments of a project's SEU in favor of pursuing existing lines of research.

A rich literature on the sunk cost effect from economics, psychology, and management suggests that people's actions are influenced by costs they have already incurred, and that these past actions will constrain decision-making in the present, even though these actions do not affect the attractiveness of available options. The intuition from both theoretical and empirical work on the topic is that people tend to use *ex post* rationalization to justify and double down on their original course of action even if it leads to continued investments in losing endeavors ("throwing good money after bad") (Arkes and Blumer, 1985; Eyster et al., 2021; Staw and Hoang, 1995; Thaler, 1999). Recent empirical examinations of the sunk cost effect indicate that it is both an intrapersonal and interpersonal effect, meaning that people will alter their choices in response to not only their *own* but also *other people's* past investments, regardless of their social closeness (Olivola, 2018).

If reviewers' perceptions of risk and a project's p(s) are influenced by a project's prior investments and funding, then indicators of the research team's *commitment* and the project's *progress* are likely to shape funding decisions and outcomes. Put differently, once prior expenditures in a project have occurred, people's estimated probability of whether a project will succeed *increases,* and results in *p(s) inflation* (Arkes and Hutzel, 2000). The effect has been linked to prospect theory, which indicates that an individual will be more likely to make a risky investment after incurring a sunk cost because they are more willing to risk small losses to obtain possible large gains in value (Kahneman and Tversky, 1979) due to loss aversion with respect to a reference point that is fixed before the costs were incurred or sunk.

Because the SEU asks reviewers to estimate a project's p(s), it is important to consider how sunk costs may alter reviewers' preferences when evaluating alternative research proposals. The above discussion on prior investments and p(s) inflation sheds some light on how reviewers may be susceptible to the sunk cost effect. It suggests that reviewers may assign higher p(s) to projects for which the research team has incurred greater prior investments. These prior investments may be correlated with project attributes, such as contingency plans, preliminary data, past purchases of expensive or specialized equipment, and continuation of existing lines of research, as well as research team attributes, such as a demonstrated track record or an experienced Principal Investigator (PI). Each of these attributes just mentioned corresponds to many of the sub-questions that F&S suggest using to arrive at a project's p(s) (i.e., see p. 20). If reviewers are susceptible to p(s) inflation due to prior investments (i.e., sunk costs), this may skew their preferences *away* from projects that propose something novel or tests something without proof of feasibility and *towards* projects that continue existing research lines—thereby increasing the difference in attractiveness (in terms of their p(s)) between the two alternative project types even further. If reviewers are affected by sunk costs in their evaluation decisions, then project evaluation outcomes are unlikely to reflect accurate judgments of a project's risk, even after accounting for their differences in value.[6]

[6] It is also possible that sunk costs may alter reviewers' perceptions of the *value* of proposed research. See Bernheim et al. (2021) and Eyster et al. (2021) for related theoretical work suggesting that people's perceptions of value may be subjective and can change over time.

Although studies on the effect of sunk costs on a project's ability to attract subsequent funding in science are rare, some recent work on the funding of grant proposals suggests that earlier funding success and greater coherence with an applicant's previous work leads to higher success in attracting subsequent funds (Ayoubi et al., 2020; Bol et al., 2018). While these papers do not investigate the reviewers' perceptions of project risk directly, it is possible that the sunk cost effect may in part be responsible for driving subsequent project funding success and investment preferences. An additional related factor to consider here is the multi-period nature of projects in science. For instance, a typical NIH R01 grant cycle lasts three years before it is up for renewal (Azoulay et al., 2011). It is likely that the sunk cost effect will be more likely to apply to multi-period projects that require large investments over time, in which the resources required at each stage are both difficult to predict ex ante and can be subject to cost increases or project delays.

Taken together, given that both risk and value are fundamental to the SEU framework, it is imperative that we systematically investigate how sunk costs may drive resource allocation decisions in science funding opportunities. An underlying goal of these studies would be to alter reviewers' reference point of prior investments in the research and to derive approaches to reframe the investment decision at each stage to be forward rather than backward looking. I conclude this section by suggesting that there are several potential study designs that could investigate whether and how sunk costs shape reviewer assessments. One plausible approach would be to blind the reviewers from information about the project team and the previous track records of the proposals. This approach is similar to the practice of double-blind reviews (e.g., Blank, 1991) but would go a step further to conceal any information about prior funding success, related work or preliminary data. Another approach to mitigate p(s) inflation is to remove known barriers to carrying out the proposed work, such as time, resource, and budget constraints. If such factors contribute meaningfully to inflating the p(s) for some projects over others, then removing them from the criteria for assessment may help level out the playing field across projects.

4. The Role of Reviewer Expertise in Assessments of Risk and Value

My final comment relates to F&S's general use of "experts" to make assessments of a project's risk and value. This stands in contrast to prior work indicating that a reviewer's expertise in the domain is likely to influence their preferences for different types of research (Li, 2017), and hence, may impact their assessments of a project's value and risk in the SEU. Although prior work suggests that one possible approach of eliciting a range of preferences is to recruit reviewers with varying degrees of intellectual distance (or knowledge overlap) to the domain (Boudreau et al., 2016), this approach may yield diminishing returns because reviewers who are too distant from the domain are unlikely to possess the relevant knowledge to make appropriate assessments of a project's risk or value (McLaughlin, 2001).

Instead, I propose that an alternative approach is to recruit reviewers with different types of expertise who can provide complementary assessments of a project's value and/or risk. Two sources of of expertise to consider are reviewers' domain knowledge and technical skills in the proposed method or technique envisaged. Whereas domain expertise refers to knowledge about the facts, principles, opinions and existing paradigms in the research area or topic, technical expertise refers to knowledge of the method, technique or strategy being proposed to carry out the research. Although these different sources of expertise may reside within the same reviewer in some fields (e.g., engineering, statistics, epidemiology) where research problems tend to involve a direct application of a modality, they might not. For example, in the field of advanced imaging research, a major challenge is that progress requires expertise in the latest advanced imaging tools and technologies and a deep understanding of the health problems to which these tools could be applied. Often, these knowledge bases are held by people with different disciplinary backgrounds (Boudreau et al., 2017; Lane et al., 2021a).

If the SEU were to select reviewers based on their pertinent knowledge base vis à vis the proposal's topic, an important question to raise is whether reviewers are likely to possess the expertise to predict the potential range of outcomes as well as make assessments of their value and probabilities of success. Instead, a more viable approach might be to develop an evaluation process whereby reviewers are asked to review a proposal focusing on either their assessments of the value or the probabilities of success associated with the revealed range of predicted outcomes. Adopting a "component" approach may be advantageous for a few reasons. First, it would reduce the cognitive load associated with multiple-criteria decision-making (Yu, 2013). Second, it may create a better match between a reviewer's expertise and the evaluation task while avoiding the potential concern of positive or negative judgments on one dimension affecting judgments in another dimension (Thorngate et al., 2010). Third, it has the potential to create alternative ways to rank proposals according to different assessment criteria, namely value, risk, and the combined SEU.

Following on this argument, another area that F&S leave as an open question is how to aggregate across reviewers' judgments using the SEU framework to come up with funding decisions. At the NIH, proposals are rank ordered, and funds are distributed in descending order until the budget is exhausted, which makes each investment a standalone decision. However, a key insight from the investment literature is that while a research proposal may appear risky as a standalone investment, this project-specific risk may be mitigated across a range of research proposals, because the risks can offset each other through diversification.

It strikes me that another potential advantage of drawing upon evaluators with different expertise dimensions is that it creates alternative approaches to rank ordering proposals—and may offer the funding agency or grant officer a way to diversify risk across the funded *portfolio* of proposals. For example, proposals can be rank ordered based on the aggregated scores among the domain experts, technical experts, or across both types of experts. Proposals can then be selected by comparing them across alternative ranking schemes. One approach might be to fund the projects that receive the highest independent rankings by the domain experts and technical experts, respectively. Another possibility might be to fund the projects that receive the highest scores across both sets of expert rankings. Ultimately, the objective of comparing the performance of these alternative ranking schemes against one another would be to offer a strategy for diversifying the portfolio of funded projects based on their relative strengths along different dimensions of assessment to improve resource allocation decisions.

In short, I believe that greater specificity into the types of experts that are selected to evaluate the proposals as well as clarity into how funding decisions will be determined are critical next steps to advancing understanding of the SEU framework.

5. Looking Ahead at Funding Risky Science: Field Experiments As a Way Forward

The SEU framework leads to a range of hypotheses about the relationships between project risk and value in the peer review process that are ripe for testing and validation. Field experiments represent a compelling methodological tool for hypothesis testing because they are useful for assessing causality in natural occurring settings as they can address the counterfactual problem of what might have happened in the absence of an intervention and can help overcome confounding issues to establish causal relationships (List and Metcalfe, 2014). To begin testing these recommendations, one potential path forward would be to partner with funding agencies and organizations to design interventions to solve problems of mutual interest. Recent experimental work conducted between the Laboratory for Innovation Science at Harvard (LISH) and the

Harvard Catalyst pilot grant program shows that partnering with grant organizations can be a promising yet pragmatic way to understand how funding allocation decisions are related to various dimensions of the peer review process, such as the evaluators' expertise in the domain area (Boudreau et al., 2016) and opinion-sharing among evaluators (Lane et al., 2021b). Over time, these experimental studies also lend themselves well to examining the implications of ex ante funding choices on intermediate- to long-run research outputs and knowledge advances via the paper trail of publications, patents, and other materials associated with the submitted proposals.

Although the findings from this recent experimental work suggests that expert reviewers may have preferences for more conservative project proposals, our understanding of the relationships between project risk (e.g., in terms of each project's novelty and feasibility), value, and funding allocation decisions remain very much in its infancy—particularly as it relates to testing alternative peer review models and approaches. Building on F&S's call for future work to test the effectiveness of the SEU framework, I advocate that we take a field experimental approach to address big picture questions, such as understanding how the SEU approach to evaluation and selection decisions systematically differs from more traditional peer review models used by large funding organizations, such as the NIH and the ERC, and also whether a portfolio view of funding opportunities, consideration of sunk costs from prior investments, or the availability of data-driven novelty and feasibility project metrics may result in the selection of projects with different risk-reward profiles. Looking forward, I believe there are many promising research questions to investigate that lie at the intersection of the peer review process, field experiments, and funding risky science.

6. Conclusion

Franzoni and Stephan offer a refreshing, alternative peer review framework to address the growing conservatism of research funding decisions. The SEU framework provides a promising approach to account for reviewers' perceptions of risk and value in their funding decisions. Like any novel approach, the SEU requires thorough testing and experimentation to improve and validate the framework. I believe that experimentation can help improve our theoretical understanding of when and under which conditions the SEU can be applied to evaluating research proposals as well as empirical understanding of whether and whose assessments of value *and* risk are informative for managing resource allocation decisions. Ultimately, there needs to be a balance between increasing the merit of evaluation processes in selecting *high quality* yet potentially risky science, and the complexity of introducing novel and time-intensive approaches to funding decisions. My comments in this response are intended to be provocative, to help us push further and deeper into demystifying the meaning of risk and uncertainty in scientific funding. I hope that these comments are helpful to Franzoni and Stephan and to other scholars interested in advancing the knowledge frontier.

Acknowledgements: I would like to thank Charles Ayoubi, Eva C. Guinan, Manuel Hoffmann, and Karim R. Lakhani for their conversations and insights that have informed this Response, as well as the constructive feedback from Research Policy editor, Ben Martin, and associate editor, Ohid Yaqub on an earlier version of this Response.

7. References

Alon, U., 2009. How to choose a good scientific problem. Molecular cell 35, 726–728.
Arkes, H.R., Blumer, C., 1985. The psychology of sunk cost. Organizational behavior and human decision processes 35, 124–140.
Arkes, H.R., Hutzel, L., 2000. The role of probability of success estimates in the sunk cost effect. Journal of Behavioral Decision Making 13, 295–306.
Ayoubi, C., Barbosu, S., Pezzoni, M., Visentin, F., 2020. What matters in funding: The value of research coherence and alignment in evaluators' decisions. Maastricht Economic and Social Research Institute on Innovation and ….

Azoulay, P., Fons-Rosen, C., Graff Zivin, J.S., 2019. Does science advance one funeral at a time? American Economic Review 109, 2889–2920.

Azoulay, P., Graff Zivin, J.S., Manso, G., 2011. Incentives and creativity: evidence from the academic life sciences. The RAND Journal of Economics 42, 527–554.

Bernheim, B.D., Braghieri, L., Martínez-Marquina, A., Zuckerman, D., 2021. A theory of chosen preferences. American Economic Review 111, 720–54.

Blank, R.M., 1991. The effects of double-blind versus single-blind reviewing: Experimental evidence from the American Economic Review. The American Economic Review 1041–1067.

Bol, T., de Vaan, M., van de Rijt, A., 2018. The Matthew effect in science funding. Proceedings of the National Academy of Sciences 115, 4887–4890.

Boudreau, K.J., Brady, T., Ganguli, I., Gaule, P., Guinan, E., Hollenberg, A., Lakhani, K.R., 2017. A field experiment on search costs and the formation of scientific collaborations. Review of Economics and Statistics 99, 565–576.

Boudreau, K.J., Guinan, E.C., Lakhani, K.R., Riedl, C., 2016. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. Management Science 62, 2765–2783.

Eyster, E., Li, S., Ridout, S., 2021. A Theory of Ex Post Rationalization. arXiv preprint arXiv:2107.07491.

Gallo, S., Thompson, L., Schmaling, K., Glisson, S., 2018. Risk evaluation in peer review of grant applications. Environment Systems and Decisions 38, 216–229.

Hackett, E.J., Chubin, D.E., 2003. Peer review for the 21st century: applications to education research. National Research Council, editor.

Jeppesen, L.B., Lakhani, K.R., 2010. Marginality and problem-solving effectiveness in broadcast search. Organization science 21, 1016–1033.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 363–391.

Lane, J., Ganguli, I., Gaule, P., Guinan, E., Lakhani, K.R., 2021a. Engineering serendipity: When does knowledge sharing lead to knowledge production? Strategic Management Journal 42.

Lane, J., Teplitskiy, M., Gray, G., Ranu, H., Menietti, M., Guinan, E., Lakhani, K.R., 2021b. Conservatism Gets Funded? A Field Experiment on the Role of Negative Information in Novel Project Evaluation. Management Science (Forthcoming).

Lee, C.J., 2015. Commensuration bias in peer review. Philosophy of Science 82, 1272–1283.

Lew, D., 2005. The Longest-Standing Math Problem. URL https://www.damninteresting.com/the-longest-standing-math-problem/ (accessed 6.12.22).

Li, D., 2017. Expertise versus Bias in Evaluation: Evidence from the NIH. American Economic Journal: Applied Economics 9, 60–92.

List, J.A., Metcalfe, R., 2014. Field experiments in the developed world: an introduction. Oxford Review of Economic Policy 30, 585–596.

Luukkonen, T., 2012. Conservatism and risk-taking in peer review: Emerging ERC practices. Research Evaluation 21, 48–60.

McLaughlin, N., 2001. Optimal marginality: Innovation and orthodoxy in Fromm's revision of psychoanalysis. Sociological Quarterly 42, 271–288.

Mount, M.P., Baer, M., Lupoli, M.J., 2021. Quantum leaps or baby steps? Expertise distance, construal level, and the propensity to invest in novel technological ideas. Strategic Management Journal 42, 1490–1515.
Mueller, J.S., Melwani, S., Goncalo, J.A., 2012. The bias against creativity: Why people desire but reject creative ideas. Psychological science 23, 13–17.
Myers, K., 2020. The elasticity of science. American Economic Journal: Applied Economics 12, 103–34.
Olivola, C.Y., 2018. The interpersonal sunk-cost effect. Psychological science 29, 1072–1083.
Page, S., 2008. The difference. Princeton University Press.
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, 2952–2957.
Staw, B.M., Hoang, H., 1995. Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. Administrative Science Quarterly 474–494.
Stokes, D.E., 1997. Pasteur's quadrant: Basic science and technological innovation. Brookings Institution Press.
Thaler, R.H., 1999. Mental accounting matters. Journal of Behavioral decision making 12, 183–206.
Thorngate, W., Dawes, R.M., Foddy, M., 2010. Judging merit. Psychology Press.
Uzzi, B., Mukherjee, S., Stringer, M., Jones, B., 2013. Atypical combinations and scientific impact. Science 342, 468–472.
Wang, J., Veugelers, R., Stephan, P., 2017. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. Research Policy 46, 1416–1436.
Yu, P.-L., 2013. Multiple-criteria decision making: concepts, techniques, and extensions. Springer Science & Business Media.