Credibility in Social and Personality Psychology
PsychOpen GOLD

Social Psychological Bulletin

Strategic Ambiguity in the Social Sciences

Willem E. Frankenhuis [1,2], Karthik Panchanathan [3], Paul E. Smaldino [4,5,6]

[1] *Department of Psychology, Utrecht University, Utrecht, the Netherlands.* **[2]** *Max Planck Institute for the Study of Crime, Security and Law, Freiburg, Germany.* **[3]** *Department of Anthropology, University of Missouri, Columbia, MO, USA.* **[4]** *Department of Cognitive and Information Sciences, University of California, Merced, CA, USA.* **[5]** *Center for Advanced Study in the Behavioral Sciences, Stanford University, Stanford, CA, USA.* **[6]** *Santa Fe Institute, Santa Fe, NM, USA.*

Social Psychological Bulletin, 2023, Vol. 18, Article e9923, https://doi.org/10.32872/spb.9923

Received: 2022-07-22 • **Accepted:** 2022-10-14 • **Published (VoR):** 2023-##-##

Handling Editors: Simine Vazire, Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Australia; Brian Nosek, University of Virginia, Charlottesville, VA, USA

Corresponding Author: Willem E. Frankenhuis, Heidelberglaan 1, 3508 TC Utrecht, The Netherlands. E-mail: w.e.frankenhuis@uu.nl

Related: This article is part of the SPB Special Topic "Is Psychology Self-Correcting? Reflections on the Credibility Revolution in Social and Personality Psychology", Guest Editors: Simine Vazire & Brian Nosek, Social Psychological Bulletin, 18, https://doi.org/10.32872/spb.v18

Abstract

In the wake of the replication crisis, there have been calls to increase the clarity and precision of theory in the social sciences. Here, we argue that the effects of these calls may be limited due to incentives favoring ambiguous theory. Intentionally or not, scientists can exploit theoretical ambiguities to make support for a claim appear stronger than it is. Practices include theory stretching, interpreting an ambiguous claim more expansively to absorb data outside of the scope of the original claim, and post-hoc precision, interpreting an ambiguous claim more narrowly so it appears more precisely aligned with the data. These practices lead to the overestimation of evidence for the original claim and create the appearance of consistent support and progressive research programs, which may in turn be rewarded by journals, funding agencies, and hiring committees. Selection for ambiguous research can occur even when scientists act in good faith. Although ambiguity might be inevitable or even useful in the early stages of theory construction, scientists should aim for increased clarity as knowledge advances. Science benefits from transparently communicating about known ambiguities. To attain transparency about ambiguity, we provide a set of recommendations for authors, reviewers, and journals. We conclude with suggestions for research on how scientists use strategic ambiguity to advance their careers and the ways in which norms, incentives, and practices favor strategic ambiguity. Our paper ends with a simple mathematical model exploring the conditions in which high-ambiguity theories are favored over low-ambiguity theories, providing a basis for future analyses.

Keywords

strategic ambiguity, theory development, formal modeling, incentive structures, theory stretching, post-hoc precision, RAPPing

Highlights

- Current incentives favor strategic ambiguity (i.e., deploying ambiguity to achieve self-interested goals).
- Researchers exploit the flexibility that theoretical ambiguity affords, intentionally or not, using practices like theory stretching and post-hoc precision, which result in the overestimation of evidentiary support for theoretical claims.
- Scientists should aim for transparency when ambiguity is unavoidable in the short run and aim for increased clarity in the long run.
- We provide recommendations for authors, editors, and funders on how to restructure incentives to disfavor strategic ambiguity and favor clarity.

"You who are so good with words
And at keeping things vague"
Joan Baez (1975), *Diamonds & Rust*

In the wake of the replication crisis, there have been numerous calls to increase the clarity and precision of theory in the social sciences. A popular solution calls for the 'formalization' of theory, expressing theory in the language of mathematics (Borsboom, 2013; Borsboom et al., 2021; Frankenhuis et al., 2013; Fried, 2020; Guest & Martin, 2021; Muthukrishna & Henrich, 2019; Navarro, 2021; Robinaugh et al., 2021; Smaldino, 2020a; van Rooij & Baggio, 2021). Formalizing theory can promote clarity and precision, for instance, by requiring the scientist to specify all constructs and their relations, make assumptions explicit, and logically deduce predictions. Of course, formal models are no panacea, and may be of limited use when they rely upon questionable assumptions. In some cases, a well-articulated verbal theory may be sufficient. The key point is that scientists should strive for more rigorous, clear, and precise theory.

Such calls for increased theoretical rigor are not new. Similar calls, in fact, predate the replication crisis (Epstein, 2008; Farrell & Lewandowsky, 2010; Gibbs, 1987; Harris, 1976), as do more general manifestos cataloguing the immature state of theory in the social sciences (Gigerenzer, 1998; Meehl, 1978; Mischel, 2008; Rozin, 2001). It is unknown whether the application of formal theory is actually increasing in response to these calls. Regardless, formal and precise theory remains rare in the social sciences—disciplines studying human behavior, including but not limited to psychology, anthropology, criminology, and sociology—though exceptions exist in several subfields. And where formalization is absent, ambiguity can flourish (Smaldino, 2017).

The persistence of ambiguous theory raises uncomfortable questions: Are social scientists aware of the problem? If so, do they see this as only a minor issue? Could it be that some researchers in some cases *prefer* ambiguous theory? A naïve explanation of the prevalence of ambiguous theory might be that producing formal and precise theory requires advanced mathematical and programming techniques that are beyond the skillset of the typical social scientist. But this cannot be the whole story. After all, social scientists routinely use equally advanced methodological and statistical techniques. Might there be structural and systemic factors—rules, norms, and institutions—limiting the spread of formal and precise theory?

The Incentives Made Me Do it

> "I have 3 recent experiences where I am publishing counter-evidence and the editor sends the ms directly to the author of the theory I am addressing and that person says, no, not counter-evidence at all because **my theory can be stretched to cover that data.**"
>
> Susan Rvachew (2021, emphasis added)

In this paper, we argue that current incentives favor *strategic ambiguity* in the development and presentation of theory. By 'strategic ambiguity' we mean the use of ambiguity to achieve self-interested goals (Eisenberg, 1984). Any statement or set of statements, including a theory or hypothesis, is 'ambiguous' to the degree that it is open to multiple interpretations (Eisenberg, 1984; Gambetta, 2011; Lee & Pinker, 2010). Like everyone else, scientists sometimes strategically deploy ambiguity for the flexibility it affords—flexibility which can be used to obscure weaknesses, deny specific interpretations, and accommodate unexpected data. In these and other ways, ambiguous theories are harder to falsify than their precise counterparts (Popper, 1963; Smaldino, 2017). Although previous research has discussed the difficulty of falsifying flexible theory (Szollosi & Donkin, 2019, 2021), there has been less research on how scientists employ ambiguity strategically in formulating theory and on the incentives that make this possible. Analyzing strategic ambiguity in communication, Eisenberg (1984) noted: "The more ambiguous the message, the greater the room for projection. When an individual projects, he or she fills in the meaning of a message in a way which is consistent with his or her own beliefs" (p. 233). In science, ambiguous theories are more prone to confirmation bias—the tendency to search for, favor, interpret, and remember information in a way that confirms or supports one's prior beliefs (Nickerson, 1998)—and are therefore more resistant to falsification when compared to clear and precise theories. As a result, the use of strategic ambiguity by scientists may come at the cost of scientific progress.

The "motte-and-bailey fallacy" (Shackel, 2005), a pernicious rhetorical trick named after a medieval castle design, illustrates the use and usefulness of strategic ambiguity.

Here, the speaker first advances an indefensible claim (the ‘bailey’). When challenged, the speaker retreats to a more modest claim (the ‘motte’) that shares some similarities with the indefensible claim. In this way, the speaker can argue that the original (and indefensible) claim has not been refuted without ever having to defend it! Worse, the speaker can accuse the challenger of being unreasonable. As Boudry and Braeckman (2011) note, “A skilled pseudoscientist switches back and forth between different versions of his theory, and may even exploit his own equivocations to accuse his critics of misrepresenting his position” (p. 150).

Ambiguity at any stage of the research process can impede scientific progress (Rohrer, 2021). For instance, during statistical hypothesis testing, ambiguity in preregistrations affords degrees of freedom that researchers might exploit by running different analyses and selectively reporting those that yield desired outcomes. One solution for removing this ambiguity is to make preregistrations machine readable (Lakens & DeBruine, 2021; van Lissa, 2022). Here, we focus on ambiguity in theory formation and communication. We build on a previous blogpost on this topic (Frankenhuis et al., 2021), an earlier book chapter (Smaldino, 2017), and a recent talk on the properties of theories that promote their proliferation (Hussey, 2022). Though we draw our examples primarily from psychology, we believe the problem of strategic ambiguity is widespread across much of the social sciences (for a recent discussion of ambiguity in criminological theory, see Niemeyer et al., 2022).

The Natural Selection of Ambiguous Theories

Scientists are, of course, motivated to discover the truth and thereby help the scientific enterprise succeed. But, being humans, scientists are also motivated to have impact and thereby help *themselves* succeed. Impact results from having ideas widely cited and discussed. One way of achieving impact is to propose theories that accurately describe known phenomena and explain and predict novel phenomena. In this case, the interests of scientists and the scientific enterprise converge. Another way is to propose theories that can be interpreted in many different ways and thereby *appear* to describe known phenomena and predict novel phenomena. The ability of ambiguous verbal models to accommodate a wide range of findings and their resistance to falsification might increase their impact on the scientific community. If the motivation for impact is strong enough, scientists might prefer the flexibility afforded by ambiguous verbal models over the rigidity imposed by precise formal models. Intentionally or not, scientists can use this flexibility to create the appearance of consistent support and progressive research programs, which may be rewarded by journals, funding agencies, hiring committees, and the press.

The problem of detecting ambiguous theories masquerading as clear ones is made worse as gatekeepers, including scholarly journals, funding agencies, and university departments, crave impact, too. These gatekeepers also benefit from ambiguous theories

https://doi.org/10.32872/spb.9923

if those theories produce research that raises a journal's impact factor, increases a funder's media visibility, or boosts a department's ranking. Intentionally or not, scholarly journals, funding agencies, and university departments may be contributing to the proliferation of ambiguous theory. And so there may be many potential levers for improving incentives. In this paper, we focus on the incentives that push scientists, rather than scientific institutions, toward strategically ambiguous theory.

We should note that our discussion of strategic ambiguity does not require *intentional* ambiguity. Although intentional ambiguity certainly occurs, incentives can favor scientists who use ambiguous theories over scientists who use precise, well-specified theories without scientists being aware of the incentive structures nor consciously choosing to be more ambiguous (Hussey, 2022; Smaldino & McElreath, 2016; Stewart & Plotkin, 2021). For example, consider that some scientists criticize the use of formal models for the unrealistic assumptions they demand. This is a valid criticism in some cases. But it is also possible that this criticism rationalizes a preference for verbal models and the ambiguity they afford. Perhaps there is also the concern that formal models will reduce the impact of—or even replace—familiar verbal models, especially a researcher's own. Some researchers may well be aware of their goals, but not the extent to which these goals end up shaping their beliefs. Our claim is that ambiguous theory proliferates especially widely in disciplines where verbal models predominate due to ambiguities in natural language, regardless of whether scientists act in good faith. We will also argue that reshaping incentives to favor transparency about ambiguity will require intentional, goal-directed action by scientists. These actions need to be supported throughout the scientific system, including scholarly journals, funding agencies, and university departments.

Outline of the Paper

We begin by arguing that there is nothing inherently wrong with theoretical ambiguity. In fact, in the early stages of theory construction, ambiguity may be inevitable—and may even be useful. But we also argue that this ambiguity should be transparent—and that scientific institutions should be designed to promote this kind of transparency. And so, in the subsequent section we provide recommendations for authors, reviewers, and journals to bring about transparency in the presentation of ambiguous theory. For instance, we provide a set of questions that can help to evaluate the extent to which a theory is ambiguous, akin to empirically-oriented checklists (Aczel et al., 2020; Flake & Fried, 2020). We conclude with suggestions for research on how scientists use strategic ambiguity to advance their careers and which norms, incentives, and practices favor strategic ambiguity. Our paper ends with a simple mathematical model exploring the conditions in which high-ambiguity theories are favored over low-ambiguity theories. This model illustrates that incentives for credit can favor low-quality, high-ambiguity science, and provides a basis for future formal analyses.

Clarity and Ambiguity Each Have Their Place

> "Preliminary operationalizations and fuzzy inferences are not a crime, but a normal starting point of scientific discovery. Yet to progress toward precise claims, the initial vagueness must be recognized and tackled in subsequent studies."
>
> Anne Scheel (2022, p. 3)

At this point, some readers might complain, "But psychology [or sociology, or anthropology, or ...] is a young science!" or "Human behavior is too complex to ever have precise theories like physics!" Depending on the details of the argument, we may be sympathetic to these complaints. There are deep and fundamental differences between physics and the social sciences (Fodor, 1974). But we think these kinds of objections, in general, miss the mark. We are not arguing that ambiguity has no place in scientific inquiry. Instead, we argue that while the transparent use of ambiguity can benefit scientific inquiry, we should not tolerate ambiguity used strategically to benefit the scientist at the expense of the scientific enterprise.

Of course, the study of human behavior faces special challenges. There is often a yawning gap between theoretical concept and empirical measurement, especially when compared to a science like physics. Physicists largely agree on how to define and measure the motion and mass of matter, while psychologists often disagree on how to define, let alone measure, the mental states and motivations of minds. Realistically, these kinds of challenges are unlikely to be resolved any time soon. Perhaps they never will. Nevertheless, at the level of theory—concepts and their relations—social scientists can and should aim for clarity.

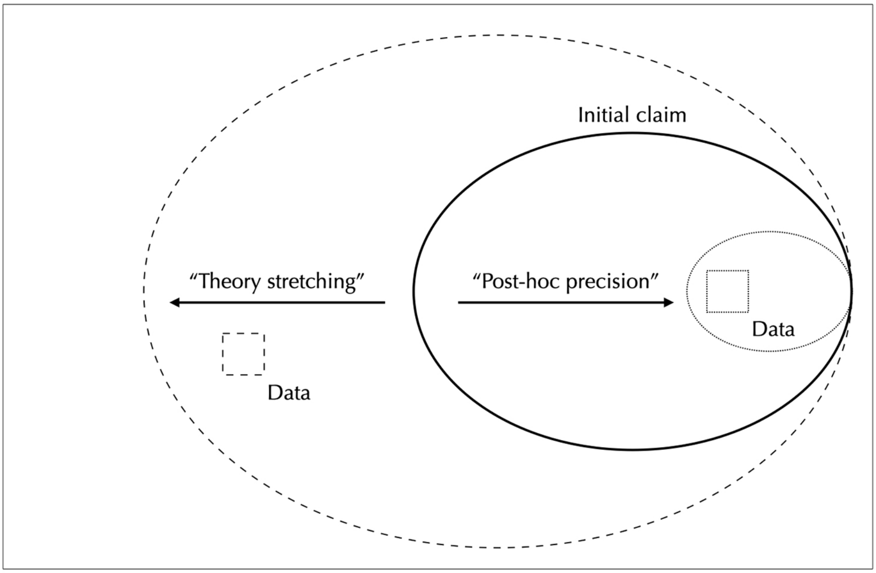
Despite its benefit to the scientific enterprise, clear theory will not settle every dispute. For instance, scientists may disagree over which empirical unit (estimated from observed data) best captures a given theoretical unit (Lundberg et al., 2021; Rohrer, 2021). Nevertheless, this kind of debate will be more productive when scientists operate within a shared framework of transparent ideas and logic, rather than a wild west world of ambiguous intuition. We highlight two specific ways in which strategically ambiguous theory can be detrimental to scientific inquiry. There are certainly others.

Theory Stretching and Post-Hoc Precision

Intentionally or not, scientists might leverage theoretical ambiguity to serve their own strategic ends. Figure 1 illustrates two different ways in which scientists might exploit the wiggle room afforded by ambiguity to make the evidence for a theoretical claim appear stronger than is warranted.

Figure 1

Illustration of how Ambiguity Affords Inferential Wiggle Room



Note. The rectangle represents a hypothetical space. The medium-sized, solid ellipse represents an initial theoretical claim. Encompassing a set of different hypotheses rather than just one, this claim is 'ambiguous' as it is open to multiple interpretations. Now, suppose data supports a hypothesis outside of the initial claim, represented by the dashed square. The large, dashed ellipse represents 'theory stretching' as the scientist swallows up data that was outside the scope of the original claim. Suppose, instead, the data supports one specific hypothesis within the scope of the original claim, represented by the dotted square. The small, dotted ellipse represents 'post-hoc precision,' in which the scientist narrows the original claim to more precisely align with the data. Both theory stretching and post-hoc precision lead to the overestimation of evidence for the original theoretical claim.

First, a scientist might engage in 'theory stretching', interpreting an ambiguous claim more expansively to absorb data outside the scope of the original claim. Susan Rvachew (2021, April 17) describes this phenomenon in the epigraph above. Theory stretching may become a repeated pattern, with each revised and expanded claim serving as the basis from which to further revise and expand, thereby swallowing up more and more data that was outside the scope of the original claim. Second, a scientist might use 'post-hoc precision', interpreting an ambiguous theoretical claim narrowly so that it appears more precisely aligned with the data. For instance, a scientist might initially, and ambiguously, claim that people living in neighborhoods characterized by resource "variability" act more impulsively. Now, suppose the data support this claim only for temporal variability (fluctuating resources across time), but not spatial variability (different city blocks having

PsychOpen GOLD

different levels of resources, which remain stable over time). They might redraw their claim to apply only to temporal variability—crucially, without being explicit about this shift.

Theory stretching and post-hoc precision both lead to the overestimation of evidence for the original claim. Note that these practices differ from HARKing—Hypothesizing After Results are Known (Kerr, 1998)—which does not start with a theoretical claim, but rather with a search for statistically significant results. One way theory stretching and post-hoc precision can be kept on a leash is when, *prior to conducting a study*, researchers make precommitments (e.g., in a Registered Report) about which specific outcomes count as support for and against a theory or hypothesis (Nosek & Errington, 2020).

There is nothing inherently wrong with theory stretching and post-hoc precision, so long as they happen transparently. After all, what is learning if not revising beliefs based on observations? But if we allow ambiguous theory to masquerade as its clear counterpart, surreptitiously shapeshifting to match the data, we end up overestimating the evidentiary support for theoretical claims. Strategic ambiguity allows researchers to observe different or even contradictory data patterns, and nevertheless interpret both patterns as being consistent with the same theory (Fried, 2020; Robinaugh et al., 2021). One particularly pernicious practice is making different theoretical claims in different papers using the same label for each claim. By making empirical evidence seem to support a claim more strongly than is warranted, strategic ambiguity distorts the scientific record. To prevent this, scientists should be transparent about their use of ambiguity, a form of intellectual humility that benefits the scientific enterprise (Bringmann et al., 2022; Frankenhuis et al., 2021; Hoekstra & Vazire, 2021).

Ambiguity Might Be Inevitable—and Perhaps Even Useful

Ambiguity might be unavoidable in the early stages of theory construction. When scientists explore new territories, they encounter novel phenomena. They may face uncertainty about how to categorize or even conceptualize these new phenomena—and yet develop concepts and explanations they must (Scheel et al., 2021). And so they present candidate explanations with coarse-grain mappings between the observed parts of the system and the parts involved in their explanation. This can become an iterative process of tinkering as it may not even be initially clear what one's hypothesis even *is* (Kauffman, 1971). It would be counterproductive to dismiss nascent theories for being ambiguous. How could they be otherwise? We cannot expect a useful map until we have charted the territory! As Paul Rozin (2001) notes, "we would do well to open our eyes more widely before we dig too deep a hole at one place in the broad and varied terrain of human social life" (p. 13).

Rozin's (2001) warning of premature excavation can be interpreted in at least two different ways, either about uncertainty reduction or about surveying the problem space. Armed with a good map of the problem space, we can reduce uncertainty without having

https://doi.org/10.32872/spb.9923

to ask ambiguous questions. Consider the way young children often play the game 20 Questions:

1. Is it Ringo Starr? *No.*
2. Is it George Harrison? *No.*
3. Is it Paul McCartney? *No.*

It would be most impressive if this strategy correctly guessed the target, but extremely unlikely as the overwhelming majority of possible solutions are not members of *The Beatles.* As adults, we know it is better to start with broad questions:

1. Is it a human? *Yes.*
2. Is this person still living? *Yes.*
3. Is this person an artist? *Yes.*

In the game of 20 Questions, broad questions can be just as precise as specific ones. In fact, the whole point of broad and precise questions is to quickly reduce the search space. Knowing that the target is a 'human' dramatically narrows the space of possible solutions. Whether broad or specific, there is little room to interpret what a question means so long as it is clear. Playing this game is similar to how Platt (1964) discussed 'strong inference' as a guide to scientific discovery, in which definitive experiments carve up the possibility space in a finer and finer grain. When the scientific terrain has already been well-mapped, definitive experiments present a powerful tool for learning about how the world works.

But what happens when we do not already know the problem space? Trying to reduce our uncertainty about a specific hypothesis would be premature. Here, we need to map the problem space. This is precisely the situation in which scientists often find themselves during the early stages of inquiry: Unclear about basic concepts and categories and unsure about the space of possible, let alone probable, hypotheses. During these early stages of scientific inquiry, ambiguity may not only be inevitable—it might actually be beneficial. Ambiguity affords more interpretive room, which might lead different scientists to think about the problem in different ways, ask different types of questions, and consider different kinds of evidence. The more diverse the pool of scientists studying the phenomenon, the more hypothetical space they are likely to cover (Hofstra et al., 2020). Some will hit upon fruitful lines of inquiry, leading others to follow. In the early stages of scientific inquiry, we should heed Daniel Dennett's (1991) warning of the 'heartbreak of premature definition,' especially when studying elusive concepts like consciousness, which remain mysterious despite sustained scientific inquiry spanning decades, if not centuries.

To be clear, we are *not* calling for the abolition of ambiguity in science. That would be both counterproductive and futile. Instead, we argue that scientists should be as clear as they can be about the ambiguities lurking in their theories, whether studying novel phenomena or mysterious ones, and everything in between. And, as we will discuss, incentives should encourage transparent ambiguity and discourage its strategic counterpart.

Towards Transparency About Ambiguity

> "Giving a bad thing a name can help to raise awareness. For example "p-hacking". What should we call it when: detailed methods & explicit models are heavily criticized while thin methods & vague verbal models pass freely? Used to see this constantly on grant panels."
>
> Richard McElreath (2021)

We have argued that social scientists should strive for theoretical clarity and be transparent about ambiguity when it is either inevitable or useful. One way to do this is to construct formal models of hypotheses and theories that clearly state their meaning and scope (Kauffman, 1971; Smaldino, 2017). Several tutorials exist to support non-modelers in developing formalizations of theories (Frankenhuis et al., 2013; Fogarty et al., 2022; Smaldino, 2020a; van Rooij & Blokpoel, 2020; Wilson & Collins, 2019). But, while modeling *is* one way to reduce ambiguity, the two issues are hardly isomorphic. So rather than rehashing advice for model-building here, we instead focus on actions that authors, reviewers, and journal editors can take to increase transparency about ambiguity in the scientific literature. Clarity should be the goal, even if some ambiguity is inevitable and perhaps useful in the early stages of theory construction. By this we mean that one should be as precise as possible, given one's current understanding, as to the *scope* of a theory or hypothesis: the set of conditions under which it is or is not expected to hold (Walker & Cohen, 1985). Clarity holds interpretative wiggle room on a leash, while ambiguity liberates it. By making assumptions, concepts and their relations, and the derivation chain from assumptions to predictions explicit, it becomes easier to decouple a theory from the scientist who proposed it, making the theory a public good for all to evaluate and use (Epstein, 2008; Guest & Martin, 2021; Meehl, 1990; Smaldino, 2017). Clarity also makes it easier for a hypothesis to be falsified, and the rate at which incorrect hypotheses are falsified influences the growth of scientific knowledge (even if this growth also depends on other factors, such as the discovery of anomalies (Kuhn, 1970) and the development of new theory that better accounts for the data (Lakatos, 1970)).

Rewarding Ambiguity and Penalizing Precision

The epigraph above points out that, all too often, grant panels, hiring and promotion committees, and journal editorial boards more harshly criticize formal and precise mod-

https://doi.org/10.32872/spb.9923

els and methods than verbal and ambiguous models and methods. Responding to McElreath's tweet, Smaldino (2020b, November 30) proposed the term 'RAPPing' to denote the practice of "rewarding ambiguity and penalizing precision". The term RAPPing is general and could be used in the context of empirical work as well—for instance, when grant proposals are criticized for the planned sample sizes they provide, whereas other proposals that do not even state their planned sample sizes sail through. Here, our focus is on theory.

Why would a scientist evaluating another's research reward theoretical ambiguity? One could argue that evaluators, being human, are imperfect, despite being motivated by good intentions. This might explain noise in evaluations, but not a bias toward ambiguity. We think there is something else going on. The mapping between theory and measurement is often inexact, especially in the social sciences. When scientists write clearly about their models and methods, this inexactness is brought into stark relief. As Julia Rohrer (2021) notes: "It is the curse of transparency that the more you disclose about your research process, the more there is to criticize" (Rohrer, 2021, December 8). Clarity can have the perverse effect of making it easier for evaluators to identify flaws that might have remained hidden in a more ambiguous description. This is especially detrimental if, as suggested by empirical research, grant reviewers weigh negative information more heavily than positive information (Teplitskiy et al., 2022). Ambiguous theory does not, of course, eliminate flaws, it merely conceals them. We hope that labelling this phenomenon draws attention to it and, hopefully, encourages remediation.

To help researchers avoid RAPPing, we list six questions evaluators can ask of any hypothesis or theory to help identify ambiguities (Table 1). For each question, we have suggested a few response options, but of course it is possible to obtain more graded estimates by rating those questions on a more finely grained scale. It is not essential that all these questions be answered in the affirmative for a hypothesis or theory to be deemed clear. For instance, while formal modeling makes it easier to satisfy questions 1–3 and 5–6, some verbal theories may be clear and precise without requiring a formal model. Furthermore, clarity is but one desirable feature of a scientific theory. However, each question answered in the negative should reduce the clarity assigned to the hypothesis or theory being evaluated. We also note that the clarity of a theory is not synonymous with its empirical testability. For example, the Hawk-Dove model has been influential in the study of human and non-human conflict but was never intended to be directly testable (Maynard Smith & Price, 1973). The Hawk-Dove model does, however, involve clear assumptions that *can* be directly compared with empirical data to assess its applicability. Thus, Table 1's question about scope directly implies questions about the specification of conditions for validation or falsification. Finally, our list is not meant to be exhaustive, but merely suggestive. Future work could explore its effectiveness and improve the items for clarifying theoretical ambiguities through empirical research. Some journals and funding agencies could even train reviewers in how to use these types of questions effectively, aiming to increase standardization in evaluations of theoretical transparency.

Table 1

Six Questions to Help Evaluate Theoretical Transparency

Question	Examples of responses
1. Is each term clearly defined?	No / some / all terms are defined
2. Are all relations between terms specified?	No / some / all relations between terms are specified
3. Are all assumptions explicitly described?	No / some / all key assumptions are discussed
4. Has the theory been formalized?	The theory does / does not have a formal basis
5. Is the scope of the theory well specified?	The conditions in which the theory does and does not apply are unstated / coarsely described / fully explicit
6. Is the theory consistent across papers?	The theory is consistent / inconsistent across papers in terms of its assumptions, scope, and predictions

Evaluating Theoretical Transparency: A Worked Example

We illustrate how asking these questions can help identify key ambiguities by considering a specific example. We use this example not because it is a unique outlier, but because it is high-profile and representative of a widespread trend across the social sciences.

In a target article in the journal *Behavioral and Brain Sciences*, Baumeister, Ainsworth, and Vohs (2016) develop an argument for the following hypothesis: “Groups will produce better results if the members are individuated than if their selves blend into the group.” Some scientists might use this hypothesis as the basis for their empirical research. However, doing so would be problematic, because as it currently stands the proposal is not falsifiable (Smaldino, 2016). Evaluating this hypothesis—as it is presented in the paper—in light of Table 1 yields an answer of “no” to every single question. The hypothesis contains undefined terms and undefined relationships; it is laden with hidden assumptions and lacks clear scope; and the hypothesis is presented as a verbal rather than a formal model. Let us interrogate the hypothesis with Questions 1 and 5.

The hypothesis takes the form of a standard conditional statement (B occurs if condition A holds). The antecedent clause, “the members are individuated [instead of] their selves [being] blend[ed] into the group,” involves some ambiguity, but for simplicity we will focus most of our attention on the consequent: “groups will produce better results.”

First, the terms are not clearly defined. To what sort of groups does the hypothesis apply? The target article discusses a very wide range of groups, including track athletes in relays, families, universities, corporations, military organizations, and governments. Does the hypothesis apply to all groups? Kindergarten classrooms? Dutch soccer players? Bob Dylan fans? Or only to groups that have particular features? If so, which features? Should these features occur in combination or is any one of them sufficient?

Second, the scope of the hypothesis is never specified. What constitutes ‘better’ results? For some groups, there is a relatively straightforward answer: a sports team performs better when they win more competitions (though even here, not all competitions are equivalent). For other groups, it is harder to pin down a single measure of success; particularly as groups often exist for multiple reasons. An intervention could increase performance on one metric while decreasing performance on another (e.g., a corporation increases its profits but employee morale decreases, leading to attrition). What happens when groups are nested within larger groups, or when individuals belong to multiple groups with non-aligned interests? These and other ambiguities of the consequent clause must be addressed before one can test whether a particular group outcome is within the scope of the theory. In its stated form, the hypothesis is suggestive and can productively drive research to answer these disambiguating questions but cannot be directly tested. It is too ambiguous to falsify.

To reiterate, ambiguity is not necessarily a bad thing; it might be inevitable or even useful in the early stages of theory construction. Not all research can or should be driven by precise hypotheses. Instead, sometimes the purpose of research must be disambiguation rather than confirmation or falsification. Transparency about a current lack of precision can and should motivate future research. However, if ambiguous theoretical and empirical work is consistently supported at the expense of more precise research proposals—if there is excessive RAPPing by grant panels, editorial boards, and hiring and promotion committees—then science is the worse for it. We need transparency about ambiguity, which helps us decide for any case whether the ambiguity is inevitable, desirable, or harmful.

Recommendations and Future Directions

> “[T]he effectiveness and the ethics of any particular communicative strategy are relative to the goals and values of the communicators in the situation.”
>
> Eric M. Eisenberg (1984, p. 238)

The authors of this paper often use formal modeling. We have often heard critics argue that formal models are not useful tools in theory construction. These critics often defend their position by arguing that formal models make assumptions that are too simplistic, too unrealistic, and too arbitrary. Instead, these critics advocate for the use of verbal models, which they argue are more complex and more complicated, and therefore more realistic. To this, we might respond by highlighting the ambiguities inherent in verbal models, including imprecise constructs, implicit assumptions, and predictions based on intuitive rather than deductive reasoning. There are, of course, merits to each side of this debate (for a historically informed discussion of the mathematization of nature, see Eronen & Romeijn, 2020). Our hope is that partisans in debates like this can find common ground in our proposal to strive for transparency in the face of ambiguity, in much the same way that proponents of more exploratory research and proponents privileging confirmatory research can find common ground in the proposal to clearly delineate these two types of research (Chambers & Tzavella, 2022; Frankenhuis & Nettle, 2018; Nosek et al., 2018; Wagenmakers et al., 2012). How can authors, reviewers, and journals promote transparency about theoretical ambiguity?

Recommendations

First and foremost, the burden should be on authors to strive for transparency about ambiguities lurking in their theories. For instance, authors could write: 'our definition [of some term] is not without problems', followed by an explanation. Or 'there is friction between one of our assumptions [about the relationship between two terms] and another assumption', followed by an explanation. If authors use ambiguity deliberately, they should signal their intentions. If they do not, they may be building a Potemkin village, presenting a façade of clarity that collapses on closer inspection—and takes down anyone who was lured into working on its construction. Nguyen (2021) argues that epistemic manipulators strategically imbue their belief system with an exaggerated sense of clarity to avoid closer inspection. Whereas a sense of confusion invites us to think more, a sense of clarity, whether real or imagined, encourages us to terminate our inquiries, protecting the belief system. If authors believe their theories are clear and precise, they should welcome scrutiny rather than assume a defensive posture. The use of transparent ambiguity is fine. It is opaque ambiguity that poses risks precisely because it can be used opportunistically, whether intentional or not. The goal should be transparency in the face of theoretical ambiguity.

Reviewers should be mindful to not penalize authors for being transparent about ambiguities—just as they should not penalize authors for being explicit about exploratory aspects of their research. Reviewers should appreciate that transparency about ambiguity is a much-needed form of intellectual humility (Hoekstra & Vazire, 2021), a move that will only benefit the scientific enterprise. And just as a culture that licenses us to be more open and explorative in empirical research can feel liberating (Frankenhuis & Nettle, 2018), so too can a culture that acknowledges and embraces transparency about ambiguity in the development and presentation of theory. In addition, reviewers may consider penalizing opaque submissions—perhaps identified using Table 1—as these compete with transparent submissions.

Journals can help to reduce the harms associated with ambiguities by encouraging the formalization of theory and transparent communication about known ambiguities (van Rooij, 2022; Jamieson & Pexman, 2020). They might consider publishing articles or special issues in which empirical researchers collaborate with modelers to formalize theories, or in which many modelers independently develop models of an influential verbal theory (van Dongen et al., 2022). In addition, journals can be explicit that they value transparency about ambiguity and encourage authors to be transparent about ambiguity and reviewers not to penalize this kind of transparency. Additionally, funding agencies can help by supporting the development of formal theory.

We have argued that theoretical ambiguity proliferates in disciplines in which verbal models predominate. Though we have strived for clarity, we are certain that some of *this* paper's content is open to multiple interpretations. We welcome criticism and hope for improved clarity as we develop and test these ideas. Toward this end, we propose two future directions, one empirical and one theoretical.

Empirical Research on Strategic Ambiguity

Empirical research could examine which norms, incentives, and practices favor the use of strategic ambiguity. This kind of research would benefit from qualitative and quantitative measures of ambiguity. A first step might be to have human reviewers evaluate the degree of ambiguity in theoretical claims, perhaps by using our Table 1. Another measure of ambiguity could be the number of distinct interpretations that different researchers provide when reading a theoretical claim. Such distinct interpretations appear to explain at least part of the variation in results obtained by the different teams involved in the various “Many Analysts” projects (Scheel, 2022). Specifically, a recent reanalysis of one “Many Analysts” study (Silberzahn et al., 2018) suggests that teams answered different versions of the underspecified research question (Auspurg & Brüderl, 2021). A more quantitative approach might involve the use of a machine learning algorithm that ‘reads’ a verbal theory and quantifies the degree of ambiguity. We are aware of one study that used Flesch's reading ease score to explore whether the readability of abstracts of scientific papers is associated with the evaluation of their quality. Specifically, in an analysis of the Research Excellence Framework, a research impact evaluation of British higher education institutions, machine learning models rated harder-to-understand abstracts better than easier-to-understand abstracts; these models had been trained using mock ratings provided by scientists (Thompson et al., 2022). This finding is consistent with incentives for strategic ambiguity. We look forward to future empirical work exploring which norms, incentives, and practices favor strategic ambiguity in high-stakes, real-world settings. Some of this work should be qualitative, investigating how grant panels, hiring and promotion committees, and journal editorial boards interpret, discuss, and evaluate formal and precise models versus verbal and ambiguous models.

More broadly, the social sciences would benefit from identifying, cataloguing, and understanding the practices of theory stretching and post-hoc precision. For instance, as has already been done for HARKing (John et al., 2012), survey studies could examine how commonly researchers self-report engaging in theory stretching and post-hoc precision (self-admission rates), what they believe the percentage of other researchers who had engaged in each behavior to be (prevalence estimate), and among those researchers who had, the percentage that would admit to having done so (admission estimate) (John et al., 2012). In addition to providing baselines, this tool could be used to compare across social science disciplines (e.g., psychology and economics) and between subfields within a discipline (e.g., cognitive and developmental psychology). Such comparisons could be used to identify factors which contribute to theoretical ambiguity. Tracking these measures over time, and measuring their responses to interventions (e.g., changes in journal policies), can provide insight into perceived theoretical progress.

Finally, we have not covered all the reasons why theoretical ambiguity may be favored. For example, another reason might be the so-called 'Guru effect' (Sperber, 2010): people judge profound that which they fail to grasp. If ambiguity leads people to feel that they do not understand, this feeling of ignorance evokes awe, and this awe increases a theory's dissemination and success, then ambiguous theory may be favored. Such a process could be studied empirically, for instance by examining whether (a) people feel that they understand ambiguous theories less well than clear ones, (b) ambiguous theories evoke more awe than clear ones, and (c) theories that evoke more awe are more likely to proliferate. Though these ideas are interesting and perhaps worth exploring, we have restricted our scope to falsifiability (rather than awe) as the mediating pathway for a theory's success.

Modeling Strategic Ambiguity: A Worked Example

We believe theoretical modeling would be useful in exploring the conditions that favor and disfavor the strategic use of ambiguity. This kind of work could follow in the tradition of treating the scientific enterprise as a cultural phenomenon and applying the logic of cultural evolutionary theory (McElreath & Smaldino, 2015; Smaldino & McElreath, 2016). Such an approach can help identify conditions in which strategic ambiguity flourishes as well as interventions that might reduce it. Of course, this kind of research would also benefit from formal definitions of strategic ambiguity in the context of scientific theory. Such definitions could build on related developments in other fields, such as political science and communication (e.g., Aragonès & Neeman, 2000; Jarzabkowski et al., 2010; Pinker et al., 2008).

As a starting point, we develop a formal model to illustrate how and why theoretical ambiguity might be favored. This is a toy model, intentionally simple, designed only for the purpose of making clear an otherwise complex point. Like all models, this model makes specific assumptions. Any conclusions drawn from this analysis only apply to scenarios in which these assumptions apply.

Suppose scientists can choose one of two strategies determining how they produce and disseminate research: A *low-ambiguity strategy* (L), in which scientists derive and test precise hypotheses from clearly specified formal models, and a *high-ambiguity* *strategy* (H), in which scientists use ambiguous language to shroud the vagueness and imprecision of their theories and hypotheses.

Scientists receive credit for their work based on value conferred by their research community. We assume that scientific inquiry involves risk, so that any individual study may fail to generate credit. A low-ambiguity researcher produces useful and repeatable results a proportion p of the time. When they do, they receive a payoff of 1. And a proportion $1-p$ of the time, the low-ambiguity researcher produces less useful results. In these cases, they receive a payoff of ε (where $\varepsilon < 1$), which without loss of generality can be set to zero. Overall, low-ambiguity researchers have an expected payoff of p. The real-world consequences of low payoffs may be severe, especially for early-career researchers, as they can result in failures to get grants, promotions, or jobs.

We assume that high-ambiguity researchers produce research that is of lower quality in terms of usefulness and repeatability when compared to research produced by low-ambiguity researchers. However, by its very nature, highly ambiguous research offers flexibility in its interpretation, meaning there is a lower risk that the researcher fails to receive credit compared with the low-ambiguity strategy. At the extreme, there is little risk when a theory can accommodate any finding and get away with it. We assume that high-ambiguity researchers receive an expected payoff of q, such that $0 < q < 1$. This payoff range captures two ideas. First, that ambiguous work is less valuable than low-ambiguity research that produces useful and precise results (i.e., $q < 1$). Second, because it can accommodate any finding, high-ambiguity research will be *perceived* as more valuable than low-ambiguity research when the latter fails to produce useful results (i.e., $q > 0$). Note that if $p < q$, the expected payoff (that is, the average payoff over a large number of studies) of high-ambiguity research is larger than the expected payoff of low-ambiguity research, even if useful and repeatable results obtained from low-ambiguity research are preferable to high-ambiguity research.

Finally, we assume that research is costly in terms of time and resources. We assign a separate cost for the low and the high-ambiguity strategy: c_L and c_H, respectively. Further, we assume that producing low-ambiguity research requires greater effort than high-ambiguity research (i.e., $c_L > c_H$).

We can use these costs and benefits to calculate the expected payoffs for each strategy. The expected payoff of a strategy is the credit received minus the time and resources spent. The expected payoff for a low-ambiguity researcher is therefore:

$$U_L = p(1) + (1-p)(0) - c_L = p - c_L \quad (1)$$

Similarly, the expected payoff for a high-ambiguity researcher is:

$$U_H = q - c_H \quad (2)$$

https://doi.org/10.32872/spb.9923

We are now able to ask when the low-ambiguity strategy is favored over the high-ambiguity strategy. This happens when $U_L > U_H$. In other words, a low-ambiguity research strategy is favored when:

$$p - c_L > q - c_H \tag{3}$$

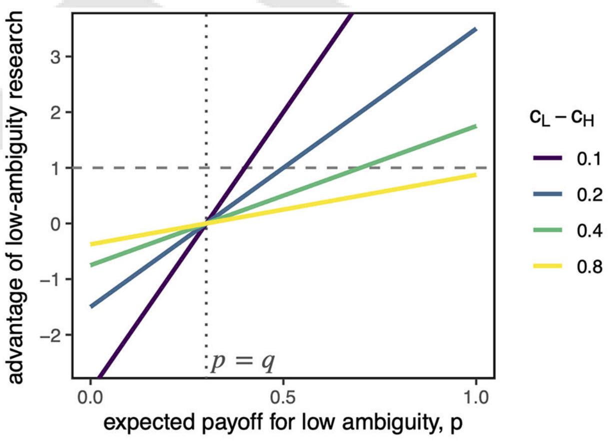
which can be rewritten as follows:

$$\frac{p - q}{c_L - c_H} > 1\,. \tag{4}$$

The left side of the inequality is the extent to which the expected credit advantage of low-ambiguity research offsets the larger cost. Two conclusions follow from this analysis. First, because the denominator is always positive, this inequality is never true if the expected credit payoff of low-ambiguity research is less than that of high-ambiguity research (i.e., $p < q$). Second, and perhaps more troubling, highly ambiguous research can be favored even if, on average, it yields a lower payoff than less ambiguous research, so long as producing the latter is sufficiently more costly. Figure 2 illustrates this relationship. Note that when low-ambiguity research is much costlier than high-ambiguity research, high-ambiguity is always incentivized.

Figure 2

Illustration of the Mathematical Model



Note. The plot shows the credit advantage to low-ambiguity research (left side of Equation 4) as a function of its expected payoff, p. The differently colored lines represent different values for the added cost to low-ambiguity research. The low-ambiguity strategy is favored only when the solid lines are above 1 on the y-axis. This requires low-ambiguity research to pay off more reliably when such research is more costly. In this example, $q = 0.3$.

Based on the logic of this simple model, we have two levers to promote greater clarity and less ambiguity. First, we should decrease the credit accruing to high-ambiguity research, perhaps through increased standards of scrutiny within a research community. Second, we should be concerned if the cost to produce low-ambiguity research in a scientific discipline is significantly higher than the cost to produce high-ambiguity research. Otherwise, low-cost, high-ambiguity research becomes incentivized and will proliferate.

Conclusion

We are confident that increased transparency about ambiguity will benefit science, but we are less clear on the value of any specific protocols for detection or enforcement. For this reason, we have avoided a more specific set of guidelines for dealing with problems related to strategic ambiguity. Systemic problems often require sustained scrutiny and systemic solutions that focus on the system in which individuals operate, rather than simply nudging individual behavior (Chater & Loewenstein, 2022). We have a long way to go. A start would be to acknowledge, as some other disciplines do (e.g., physics, biology), that developing theory is not something that empirical researchers can just do 'on the side', but rather is a professional skill that requires specialized training. Where are the jobs for theoreticians in our field? Where is funding for them? In our view, such factors contribute to the current incentives not favoring rigorous theory. Ideally, proposed solutions operate in concert, targeting multiple components of a system, as is done by funder and journal partnerships rewarding transparency in empirical research (Chambers & Tzavella, 2022; Munafò, 2017). The challenge of how to improve incentives for better theory in the social sciences is underexplored. We hope metascience researchers will make progress on this challenge in the coming years.

Funding: WEF's contributions have been supported by the Dutch Research Council (V1.Vidi.195.130) and the James S. McDonnell Foundation (https://doi.org/10.37717/220020502).

Acknowledgments: We thank Balazs Aczel, Tasha Fairfield, Sander Koole, Daniel Nettle, Glenn Roisman, Anne Scheel, Sven Arend Ulpts, Stefan Vermeent, Ethan Young, and one anonymous reviewer for feedback on previous versions of this manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Author Contributions: *Willem E. Frankenhuis*—Idea, conceptualization | Writing | Feedback, revisions | Visualization (data presentation, figures, etc.) | Supervision, mentoring | Project coordination, administration. *Karthik Panchanathan*—Idea, conceptualization | Writing | Feedback, revisions | Visualization (data presentation, figures, etc.). *Paul E. Smaldino*—Idea, conceptualization | Writing | Data analysis | Feedback, revisions | Visualization (data presentation, figures, etc.).

References

Aczel, B., Szaszi, B., Sarafoglou, A., Kekecs, Z., Kucharský, Š., Benjamin, D., Chambers, C. D., Fisher, A., Gelman, A., Gernsbacher, M. A., Ioannidis, J. P., Johnson, E., Jonas, K., Kousta, S., Lilienfeld, S. O., Lindsay, D. S., Morey, C. C., Munafò, M., Newell, B. R., & Wagenmakers, E.-J. (2020). A consensus-based transparency checklist for social and behavioral researchers. *Nature Human Behaviour, 4*, 4–6. https://doi.org/10.1038/s41562-019-0772-6

Aragonès, E., & Neeman, Z. (2000). Strategic ambiguity in electoral competition. *Journal of Theoretical Politics, 12*(2), 183–204. https://doi.org/10.1177/0951692800012002003

Auspurg, K., & Brüderl, J. (2021). Has the credibility of the social sciences been credibly destroyed? Reanalyzing the "many analysts one data set" project. *Socius: Sociological Research for a Dynamic World, 7*, 1–14. https://doi.org/10.1177/23780231211024421

Baumeister, R. F., Ainsworth, S. E., & Vohs, K. D. (2016). Are groups more or less than the sum of their members? The moderating role of individual identification. *Behavioral and Brain Sciences, 39*, Article e137. https://doi.org/10.1017/S0140525X15000618

Baez, J. (1975). Diamonds & Rust [Song]. On *Diamonds & Rust.* A&M Records.

Borsboom, D. (2013, November 20). *Theoretical amnesia.* Center for Open Science. Retrieved on April 7th 2022, from http://osc.centerforopenscience.org/2013/11/20/theoretical-amnesia/

Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science, 16*(4), 756–766. https://doi.org/10.1177/1745691620969647

Boudry, M., & Braeckman, J. (2011). Immunizing strategies and epistemic defense mechanisms. *Philosophia, 39*, 145–161. https://doi.org/10.1007/s11406-010-9254-9

Bringmann, L. F., Elmer, T., & Eronen, M. I. (2022). Back to basics: The importance of conceptual clarification in psychological science. *Current Directions in Psychological Science.* Advance online publication. https://doi.org/10.1177/09637214221096485

Chambers, C. D., & Tzavella, L. (2022). The past, present and future of registered reports. *Nature Human Behaviour, 6*, 29–42. https://doi.org/10.1038/s41562-021-01193-7

Chater, N., & Loewenstein, G. (2022, March 1). The i-frame and the s-frame: How focusing on the individual-level solutions has led behavioral public policy astray. *Behavioral and Brain Sciences.* Advance online publication. https://doi.org/10.1017/S0140525X22002023

Dennett, D. C. (1991). *Consciousness explained.* Little, Brown & Co.

Eisenberg, E. M. (1984). Ambiguity as strategy in organizational communication. *Communication Monographs, 51*(3), 227–242. https://doi.org/10.1080/03637758409390197

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation, 11*(4), Article 12. http://jasss.soc.surrey.ac.uk/11/4/12.html

Eronen, M. I., & Romeijn, J. W. (2020). Philosophy of science and the formalization of psychological theory. *Theory & Psychology, 30*(6), 786–799. https://doi.org/10.1177/0959354320969876

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science, 19*(5), 329–335. https://doi.org/10.1177/0963721410386677

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science, 3*(4), 456–465. https://doi.org/10.1177/2515245920952393

Fodor, J. (1974). The disunity of science as a working hypothesis. *Synthese, 28*, 97–115. https://doi.org/10.1007/BF00485230

Fogarty, L., Ammar, M., Holding, T., Powell, A., & Kandler, A. (2022). Ten simple rules for principled simulation modelling. *PLoS Computational Biology, 18*(3), Article e1009917. https://doi.org/10.1371/journal.pcbi.1009917

Frankenhuis, W. E., & Nettle, D. (2018). Open science is liberating and can foster creativity. *Perspectives on Psychological Science, 13*(4), 439–447. https://doi.org/10.1177/1745691618767878

Frankenhuis, W. E., Panchanathan, K., & Barrett, H. C. (2013). Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Developmental Science, 16*(4), 584–598. https://doi.org/10.1111/desc.12053

Frankenhuis, W. E., Panchanathan, K., & Smaldino, P. (2021, May 12). Strategic ambiguity or unknown unknowns? [Blog post]. https://leonidtiokhin.medium.com/strategic-ambiguity-or-unknown-unknowns-fe1596e887f3

Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry, 31*(4), 271–288. https://doi.org/10.1080/1047840X.2020.1853461

Gambetta, D. (2011). *Codes of the underworld.* Princeton University Press.

Gibbs, J. P. (1987). The state of criminological theory. *Criminology, 25*(4), 821–840. https://doi.org/10.1111/j.1745-9125.1987.tb00821.x

Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology, 8*(2), 195–204. https://doi.org/10.1177/0959354398082006

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science, 16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Harris, R. J. (1976). The uncertain connection between verbal theories and research hypotheses in social psychology. *Journal of Experimental Social Psychology, 12*(2), 210–219. https://doi.org/10.1016/0022-1031(76)90071-8

Hoekstra, R., & Vazire, S. (2021). Aspiring to greater intellectual humility in science. *Nature Human Behaviour, 5*, 1602–1607. https://doi.org/10.1038/s41562-021-01203-8

Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences of the United States of America, 117*(17), 9284–9291. https://doi.org/10.1073/pnas.1915378117

Hussey, I. (2022, February 14). *The best theory is a flawed one: lessons from implicit bias research* | RIOTS Club [Video]. YouTube. Retrieved on April 6th 2022, from https://ww.youtube.com/watch?v=GvZO_Xy5SdM

Jamieson, R. K., & Pexman, P. M. (2020). Moving beyond 20 questions: We (still) need stronger psychological theory. *Canadian Psychology, 61*(4), 273–280. https://doi.org/10.1037/cap0000223

Jarzabkowski, P., Sillince, J. A., & Shaw, D. (2010). Strategic ambiguity as a rhetorical resource for enabling multiple interests. *Human Relations, 63*(2), 219–248. https://doi.org/10.1177/0018726709337040

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Kauffman, S. A. (1971). Articulation of parts explanations in biology and the rational search for them. *Boston Studies in the Philosophy of Science, 8*, 257–272. https://doi.org/10.1007/978-94-010-3142-4_18

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). Cambridge University Press.

Lakens, D., & DeBruine, L. M. (2021). Improving transparency, falsifiability, and rigor by making hypothesis tests machine-readable. *Advances in Methods and Practices in Psychological Science, 4*(2), 1–12. https://doi.org/10.1177/2515245920970949

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological Review, 117*(3), 785–807. https://doi.org/10.1037/a0019688

Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review, 86*(3), 532–565. https://doi.org/10.1177/00031224211004187

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature, 246*, 15–18. https://doi.org/10.1038/246015a0

McElreath, R. [@rlmcelreath]. (2021, November 16). *Giving a bad thing a name can help to raise awareness. For example "p-hacking". What should we call it when* [Tweet]. Twitter. Retrieved on April 10th 2022, from https://twitter.com/rlmcelreath/status/1328267803384836096

McElreath, R., & Smaldino, P. E. (2015). Replication, communication, and the population dynamics of scientific discovery. *PLOS ONE, 10*(8), Article e0136088. https://doi.org/10.1371/journal.pone.0136088

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46*(4), 806–834. https://doi.org/10.1037/0022-006X.46.4.806

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*(1), 195–244. https://doi.org/10.2466/pr0.1990.66.1.195

Mischel, W. (2008, December 1). The toothbrush problem. *APS Observer, 21*(11). https://www.psychologicalscience.org/observer/the-toothbrush-problem

Munafò, M. R. (2017). Improving the efficiency of grant and journal peer review: Registered reports funding. *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco, 19*(7), 773–773. https://doi.org/10.1093/ntr/ntx081

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour, 3*, 221–229. https://doi.org/10.1038/s41562-018-0522-1

Navarro, D. J. (2021). If mathematical psychology did not exist we might need to invent it: A comment on theory building in psychology. *Perspectives on Psychological Science, 16*(4), 707–716. https://doi.org/10.1177/1745691620974769

Nguyen, C. T. (2021). The seductions of clarity. *Royal Institute of Philosophy Supplements, 89*, 227–255. https://doi.org/10.1017/S1358246121000035

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Niemeyer, R. E., Proctor, K. R., Schwartz, J., & Niemeyer, R. G. (2022). *Are most published criminological research findings wrong? Taking stock of criminological research using a Bayesian simulation approach.* OSF Preprints. https://doi.org/10.31219/osf.io/mhv8f

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America, 115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Nosek, B. A., & Errington, T. M. (2020). The best time to argue about what a replication means? Before you do it. *Nature, 583*, 518–520. https://doi.org/10.1038/d41586-020-02142-6

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of Sciences of the United States of America, 105*(3), 833–838. https://doi.org/10.1073/pnas.0707192105

Platt, J. R. (1964). Strong inference. *Science, 146*(3642), 347–353. https://doi.org/10.1126/science.146.3642.347

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge.* Routledge.

Robinaugh, D. J., Haslbeck, J. M. B., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspectives on Psychological Science, 16*(4), . 725–743. https://doi.org/10.1177/1745691620974697

Rohrer, J. (2021, December 8). *Who would win, 100 duck-sized strategic ambiguities vs. 1 horse-sized structured abstract?* [Blog post]. http://www.the100.ci/2021/12/08/who-would-win-100-duck-sized-strategic-ambiguities-vs-1-horse-sized-structured-abstract/

Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review, 5*(1), 2–14. https://doi.org/10.1207/S15327957PSPR0501_1

Rvachew, S. [@ProfRvach]. (2021, April 17). *I have 3 recent experiences where I am publishing counter-evidence and the editor sends the ms directly to the author* [Tweet]. Twitter. Retrieved on April 10th 2022, from https://twitter.com/ProfRvach/status/1383390130946207746

Shackel, N. (2005). The vacuity of postmodernist methodology. *Metaphilosophy, 36*(3), 295–320. https://doi.org/10.1111/j.1467-9973.2005.00370.x

Scheel, A. M. (2022). Why most psychological research findings are not even wrong. *Infant and Child Development, 31*(1), Article e2295. https://doi.org/10.1002/icd.2295

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why hypothesis testers should spend less time testing hypotheses. *Perspectives on Psychological Science, 16*(4), 744–755. https://doi.org/10.1177/1745691620966795

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., & Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science, 1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Smaldino, P. E. (2016). Not even wrong: Imprecision perpetuates the illusion of understanding at the cost of actual understanding. *Behavioral and Brain Sciences, 39*, Article e163. https://doi.org/10.1017/S0140525X1500151X

Smaldino, P. E. (2017). Models are stupid, and we need more of them. In R. R. Vallacher, S. J. Read, & A. Nowak (Eds.), *Computational social psychology* (pp. 311–331). Routledge.

Smaldino, P. E. (2020a). How to translate a verbal theory into a formal model. *Social Psychology, 51*(4), 207–218. https://doi.org/10.1027/1864-9335/a000425

Smaldino, P. E. [@psmaldino]. (2020b, November 30). Rewarding Ambiguity, Penalizing Precision — or RAPPing. [Tweet]. Twitter. https://twitter.com/psmaldino/status/1333663286395432964?s=20&t=I3WMy0o_Kggu4oSfMKjyYg

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science, 3*(9), Article 160384. https://doi.org/10.1098/rsos.160384

Sperber, D. (2010). The guru effect. *Review of Philosophy and Psychology, 1*, 583–592. https://doi.org/10.1007/s13164-010-0025-0

Stewart, A. J., & Plotkin, J. B. (2021). The natural selection of good science. *Nature Human Behaviour, 5*, 1510–1518. https://doi.org/10.1038/s41562-021-01111-x

Szollosi, A., & Donkin, C. (2019). Neglected sources of flexibility in psychological theories: From replicability to good explanations. *Computational Brain & Behavior, 2*, 190–192. https://doi.org/10.1007/s42113-019-00045-y

Szollosi, A., & Donkin, C. (2021). Arrested theory development: The misguided distinction between exploratory and confirmatory research. *Perspectives on Psychological Science, 16*(4), 717–724. https://doi.org/10.1177/1745691620966796

Thompson, J., Dreber, A., Gaunt, T. R., Gordon, M., Holzmeister, F., Huber, J., Johannesson, M., Kirchler, M., Lyon, M., Penton-Voak, I., Pfeiffer, T., & Munafò, M. R. (2022, April 10). Using prediction markets to estimate ratings of academic research quality in a mock Research Excellence Framework exercise. https://doi.org/10.31222/osf.io/gsc8f

Teplitskiy, M., Peng, H., Blasco, A., & Lakhani, K. R. (2022). Is novel research worth doing? Evidence from peer review at 49 journals. *Proceedings of the National Academy of Sciences of the United States of America, 119*(47), Article e2118046119. https://doi.org/10.1073/pnas.2118046119

van Dongen, N. N. N., Finnemann, A., de Ron, J., Tiokhin, L., Wang, S. B., Algermissen, J., Altmann, E. C., Chuang, L., Dumbravă, A., Bahník, Š., Fuenderich, J., Geiger, S. J., Gerasimova, D., Golan, A., Herbers, J., Jekel, M., Lin, Y., Moreau, D., Oberholzer, Y., . . . Borsboom, D. (2022, August 24). Many modelers. *PsyArXiv*. https://doi.org/10.31234/osf.io/r5yfz

van Lissa, C. J. (2022). Complementing preregistered confirmatory analyses with rigorous, reproducible exploration using machine learning. *Religion, Brain & Behavior*. Advance online publication. https://doi.org/10.1080/2153599X.2022.2070254

van Rooij, I. (2022). Psychological models and their distractors. *Nature Reviews Psychology, 1*(3), 127–128. https://doi.org/10.1038/s44159-022-00031-5

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science, 16*(4), 682–697. https://doi.org/10.1177/1745691620970604

van Rooij, I., & Blokpoel, M. (2020). Formalizing verbal theories: A tutorial by dialogue. *Social Psychology, 51*(5), 285–298. https://doi.org/10.1027/1864-9335/a000428

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. J. L., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*(6), 632–638. https://doi.org/10.1177/1745691612463078

Walker, H. A., & Cohen, B. P. (1985). Scope statements: Imperatives for evaluating theory. *American Sociological Review, 50*(3), 288–301. https://doi.org/10.2307/2095540

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife, 8*, Article e49547. https://doi.org/10.7554/eLife.49547

PSPS Polish Social Psychological Society

Social Psychological Bulletin (SPB) is an official journal of the Polish Social Psychological Society (PSPS).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.