PsychOpen GOLD

Social Psychological Bulletin

Psychology's Reform Movement Needs a Reconceptualization of Scientific Expertise

Duygu Uygun Tunç[1,2§], Mehmet Necip Tunç[3§]

[1] *Philosophy & Ethics Group, Eindhoven University of Technology, Eindhoven, the Netherlands.* **[2]** *Philosophy Department, Middle East Technical University, Ankara, Turkey.* **[3]** *Social Psychology Department, Tilburg University, Tilburg, the Netherlands.*

[§]*These authors contributed equally to this work.*

Social Psychological Bulletin, 2023, Vol. 18, Article e10303, https://doi.org/10.32872/spb.10303

Received: 2022-09-15 • **Accepted:** 2023-02-18 • **Published (VoR):** 2023-11-17

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

Corresponding Author: Duygu Uygun Tunç, Eindhoven University of Technology, Atlas 9.327, 5600 MB Eindhoven, the Netherlands. E-mail: duygu.uygun@outlook.com

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

Science is supposed to be a self-correcting endeavor, but who is "the scientific expert" that corrects faulty science? We grouped traditional conceptualizations of expertise in psychology under three classes (substantialist, implicitist, and social conventionalist), and then examined how these approaches affect scientific self-correction in reference to various components of the credibility crisis such as fraud/QRPs, the inadequate number of replication studies, challenges facing big team science, and perverse incentives. Our investigation pointed out several problems with the traditional views. First, traditional views conceptualize expertise as something possessed, not performed, ignoring the epistemic responsibility of experts. Second, expertise is conceived as an exclusively individual quality, which contradicts the socially distributed nature of scientific inquiry. Third, some aspects of expertise are taken to be implicit or relative to the established research practices in a field, which leads to disputes over replicability and makes it difficult to criticize mindless scientific rituals. Lastly, a conflation of expertise with eminence in practice creates an incentive structure that undermines the goal of self-correction in science. We suggest, instead, that we conceive an expert as a reliable informant. Following the extended virtue account of expertise, we propose a non-individualist and a performance-based model, and discuss why it

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

does not suffer from the same problems as traditional approaches, and why it is more compatible with the reform movement's goal of creating a credible psychological science through self-correction.

Keywords

expert, epistemic responsibility, extended cognition, distributed cognition, self-correction, virtue epistemology

Highlights

- We examine how the established research tradition in psychology conceives scientific expertise, its philosophical/sociological roots, and how it hampers reform efforts to increase psychology's self-correction capacity.
- We propose a novel performance-based and non-individualist conception of expertise in terms of informant reliability.
- We discuss how conceiving expertise as informant reliability will help to devise better science policy to increase self-correction and to better understand and counter some of the reactions to scientific reform.

Psychology's credibility crisis showed that science may dramatically fail to accomplish self-correction, at least for an extended period of time. In our opinion, one reason for this may be that the agent who is supposed to implement self-correction in science, that is, the scientific expert, is not always identified and evaluated with a view to facilitating scientific self-correction. The scientific expert is the main agent of scientific inquiry, and how scientific expertise is conceptualized has direct consequences for how the scientific record is produced and corrected. Many important normative judgments regarding scientific outcomes such as "quality," "value" or "excellence" are also conceived with direct reference to experts. We also observe that the very conceptualization of expertise has an important role in shaping the appeal of science policy suggestions by psychology's reform movement (see, e.g., Spellman, 2015; Vazire, 2018). Although the very definition of scientific expertise has not been a topic of reflection itself, a majority of these science policy proposals, and open science principles in general, imply a need for a novel perspective on scientific expertise; namely, a non-individualist model of expertise that focuses on competent and responsible epistemic performance rather than the possession of eminent credentials or unique knowledge and skills. There are substantial parallels between the conceptual implications of the reform proposals and the extended virtue account of expertise (Uygun Tunç, 2022). The extended virtue account defines an expert as a reliable informant in a particular scientific domain, which means they are a *competent* and *responsible* source for answers to questions in that domain. We also do not conceive experts as necessarily individual entities. Depending on the research question at hand, a sufficiently reliable informant may have to be identified in a supra-individual

PsychOpen GOLD

(i.e., socially extended) agent, because serving as a reliable informant may in some situations require fulfilling multi-faceted, socially distributed cognitive tasks to produce scientific knowledge (such as when dealing with complex interdisciplinary questions), and forming large research collaborations with other scientists for this purpose. In such cases, the whole research team should then be regarded as the relevant scientific expert (with all the related credit and accountability), because only it can deliver the relevant expert performance.

However, the mainstream understanding of expertise in psychological science is far from this conception. It relies on criteria for defining and evaluating expertise that are incompatible with the non-individualist, epistemic performance-focused model of expertise that is indirectly but widely implied in the scientific reform context. These criteria for expertise arguably derive from some of the more traditional or widely endorsed theoretical perspectives on expertise, which also shape the mainstream understanding of expertise in psychological science. We can roughly categorize these as the substantialist, implicitist, and the social conventionalist views of expertise. Most importantly, we argue that the prevalent endorsement of criteria for expertise that derive from these older theoretical perspectives currently hampers the reform movement by leading to resistance to reform-minded ideas on topics such as the need for and the value of direct replications, big team science, fraud/ questionable research practices, and perverse incentives.

Traditional Perspectives on Expertise

Since research on expertise is a widely relevant and transdisciplinary topic, there are not only different accounts of expertise, but also different kinds of accounts. We can very roughly distinguish three kinds of accounts of expertise: substantialist, implicitist, and social conventionalist. These categories are far from being exclusive and are only meant to capture the most salient characteristics of different accounts of expertise as they directly apply to the research tradition of psychology, and thus any particular account in the vast, multidisciplinary expertise literature may feature multiple aspects or dimensions of expertise.

1. Substantialist view: We can refer to those accounts of expertise that focus on certain substantial properties of individuals as substantialist. From this perspective, an expert is someone who has more knowledge, intelligence, experience, or skills in a domain than most people. The substantialist view has a strong affinity with the way we common-sensically or traditionally understand and attribute expertise. The dictionary definitions of expertise are often substantialist. Merriam-Webster Dictionary (n.d.) defines an expert as "one with the special skill or knowledge representing mastery of a particular subject." In the Cambridge Dictionary (n.d.), an expert is defined as "a person with a high level of knowledge or skill relating to a particular subject or activity." According to the Oxford English Dictionary (Oxford

PsychOpen GOLD

University Press, n.d.), 'expert' refers to persons who "gained skill from experience" or possess "special knowledge or skill." The substantialist view is not restricted to dictionaries. More technical definitions in philosophy are also highly similar (e.g., Brewer, 1998; Coady, 2012; Fricker, 2006; Goldman, 2001; also Watson, 2022). Most famously, Goldman (2001) calls this view "objective expertise" and defines an expert as someone who possesses a depth of knowledge and experience in a domain, and the right kind of skills for applying this knowledge to new questions in that domain. In the substantialist sense, scientific expertise is something to be *had* rather than *done*. It consists in a set of "epistemic possessions" and these categorically distinguish the owner from others (lay) in a way that justifies their epistemic authority, as long as and to the extent that they maintain their special knowledge and skills.

2. Implicitist view: This view of expertise is a version of the substantialist view, but in contrast it takes the most important epistemic possessions of experts to have an inexplicable and esoteric character (Collins, 2018; Collins & Evans, 2008), such as implicit knowledge about how to best implement a research procedure—also called tacit knowledge (Polanyi, 1966).[1] Possessing scientific expertise in this sense similarly distinguishes one from others and justifies their status as an epistemic authority, which they indefinitely maintain by virtue of possessing their unique, non-transferable and unmeasurable knowledge and skills. But unlike transferable and measurable knowledge and skills, tacit knowledge cannot be confirmed or criticized by others who are not sufficiently initiated into or internalized the relevant scientific practices, hence their esoteric character. As Collins and Evans (2008, p. 6) put it, tacit knowledge is "the deep understanding one can only gain through social immersion in groups who possess it". For this reason, the implicit view has some affinity with the social conventionalist view of expertise (below) that the substantialist view lacks. Namely, if expertise necessarily requires tacit knowledge and tacit knowledge can only be acquired through immersion in certain communities of practice, then expertise consists of social group membership in some important sense. The only possible indicator of expertise in this sense can be experience, although the proponents of this view (see, e.g., Collins & Evans, 2008) maintain that tacit knowledge cannot be reliably assessed by outsiders or the less initiated to the relevant practices.
3. Social conventionalist view: In sociology and related fields, expertise tends to be understood more as an ascribed social role than individual possession of special knowledge or skill. In this sense, expertise is akin to a label or badge. For instance, Martin (1973) maintains that "[l]ike all statuses, that of 'expert' is a social artifact.

1) Tacit knowledge comprises both knowledge and skills, because a skill is a species of knowledge–knowledge-how or practical knowledge. See also Polanyi (1966).

PsychOpen GOLD

Expertness is an ascribed quality, a badge." According to Agnew et al. (1997), the minimum criterion of expertise is the existence of a large enough social group who considers someone an expert. They explicitly emphasize the contrast between this social conventionalist view and the substantialist view: "Expertise is not synonymous with knowledge. Expertise, unlike knowledge, does not reside in the individual…Whether or not an individual is selected to serve in an expert role for a constituency is often independent of the absolute accuracy of their knowledge" (for more elaboration on this point, see Quast, 2018). The most common criterion for expertise in this sense is eminence or reputation, usually assessed on the basis of one's credentials, such as education, affiliations, publications in respected venues, as well as collection of awards and grants. This criterion for expertise implies that the epistemic authority of experts does not necessarily require objective justification. Accordingly, the epistemologist Goldman (2001) calls this view "reputational expertise," while stressing in contrast the primacy of substantive or "objective" expertise; namely that being taken as an expert by a group of people does not necessarily make one objectively an expert. Collins and Evans (2008) similarly say that "[t]o treat expertise as real and substantive is to treat it as something other than relational. Relational approaches take expertise to be a matter of experts' relations with others. The notion that expertise is only an 'attribution'—the, often retrospective, assignment of a label—is an example of a relational theory."[2]

Problems of the Traditional Views of Expertise

Several problems can be identified with these perspectives, particularly in terms of their implications for the credibility reform movement. Firstly, expertise refers to the *possession* of relevant knowledge and skills, rather than a *performance*. This is problematic from a reform-minded perspective, because possessing these does not automatically yield the competent and responsible performance which we require of scientific experts. One may possess all the pertinent skills and knowledge, but putting these to work consistently with a view to truly inform and to avoid misinforming is an epistemic performance that implies further qualities, such as being suitably motivated. Therefore, knowledge and skills should be necessary but not sufficient for being an expert.

The substantialist view of expertise (as well as its implicitist variant), with its emphasis on the possession of knowledge/skills as the defining character of an expert, says virtually nothing about the *responsibility* of an expert as an informant in reference to the scientific literature and, thereby, for instance, disconnects QRPs/fraud from the conceptualization of expertise. The implicitist view further complicates this problem by declaring the most important aspects of expertise to be inscrutable, thereby undermining

2) On the other hand, Collins and Evans argue that substantive expertise consists in certain practices of expert groups, to which individuals socialize to become experts (*viz.* implicitism).

PsychOpen GOLD

individual accountability in dealing with disputes over replication failures. The social conventionalist view similarly lacks any reference to the responsibility of an expert, and due to the central value it attaches to eminence, it sometimes functions in practice as a barrier against accountability in cases of fraud/questionable research practices and motivates mindless statistical/methodological rituals.

Secondly, focusing on epistemic possessions rather than performance makes one's conception of expertise necessarily individualistic. The substantialist view (and its implicitist variant) conceives expert competence as a psychological/cognitive property, hence makes it very difficult to conceive group experts, as if a group expert must imply a mysterious entity like a group mind. However, contemporary scientific problems are too complex to be handled effectively by single individuals. In order to demonstrate a sufficiently competent epistemic performance as required by a complex research question, scientists must often integrate their skills with those of other scientists (and with technological tools) in a way that brings about a higher degree of competence. Thus, we need a concept of expertise that does not necessarily commit to individuals but can be applied broadly to all competent providers of information in a scientific domain, whether they be large collectives, smaller research teams, or individual agents.

Thirdly, the established research tradition of psychology takes some of the most important qualities of an expert to be uniquely personal, such as having a gifted intuition for a good experimental design or a successful hypothesis. Academic training can equip someone with certain necessary skills, but what makes one a truly successful scientist is not taught or explained—it is either in you or not. A famous (or notorious) example is the reference by Baumeister (2016) to the "intuitive flair" of a good experimentalist. Further, there is a widespread notion that one learns the "arts of the trade" by immersion in the scientific community. Some research practices hardly find rational justification in textbooks or through critical reflection, which risks turning them into scientific rituals.

Fourthly, the substantialist understanding of expertise is wedded (not always free of tension) to a social conventionalist one: Although the sense of the word expertise is commonly understood in line with the substantialist view in psychology, in practice the conferral of expert status and the evaluation of expertise is based chiefly on eminence and reputation (with the hope or wishful thinking that these will strongly correlate). Actual knowledge and skill are difficult to identify directly, while relying on academic renown is naturally easy. Most people consider possession of the right epistemic qualifications meaningless unless it is coupled with recognition from the scientific community; on the other hand, we see that social recognition can practically suffice for an expert status even in the absence of the right epistemic qualifications/motivations.

The Extended Virtue Model of Expertise

The reform movement needs a conceptualization of expertise that does not have these shortcomings. The *extended virtue model* of scientific expertise (see Uygun Tunç, 2022)

PsychOpen GOLD

can be a viable alternative, and it already strongly resonates with the social organization of science that is envisioned by psychology's reform movement. The extended virtue model takes *informant reliability* as the core of expertise and accordingly defines a scientific expert as a *reliable informant* in a particular scientific domain. This definition centers on the social epistemic *function* of expertise—a person is a scientific expert only to the extent that they consistently deliver a competent and responsible performance of informing others and effectively avoiding misinforming.

There are several functionalist accounts of expertise in the philosophical and sociological literature, more recent than the traditional perspectives we mentioned, and these focus on one's performance or function of delivering specialized information. For instance, Mieg (2006) maintains defines expertise as a connection that obtains between a person and a function. Quast (2018) characterizes expertise as a special service relation between a provider and a client.[3] The most common indicator of expertise in this sense is one's track record. The extended virtue account bears a degree of similarity to these functionalist views of expertise by virtue of emphasizing performance as an informant and its indicators such as an expert's track record. But we refrain from reducing expertise to a service function, which must always be defined in reference to the demands of a clientele. Scientific experts do not have to pursue research questions that actually interest certain social groups, policy makers or science funders. A significant portion of basic scientific research aims to answer questions that are posed by no social actor in particular. These may stem from theoretical puzzles, gaps in scientific knowledge, intellectual curiosity, or explore uncharted territories. The relevant social epistemic function of a scientific expert is thus to expand and solidify our common knowledge base. Further, the existing formulations of the functionalist view are similarly individualist as the substantialist view (although not necessarily so), which the extended virtue account avoids. We will not go into a detailed formulation of the extended virtue account of expertise here, as one of us has done elsewhere (Uygun Tunç, 2022), but present a brief outline in the following to explain what we understand from *informant reliability*.

Two Components of Informant Reliability

What does informant reliability in a scientific domain comprise? We analyze informant reliability in terms of *competence* and *responsibility*. By comprising competence as one of the two components of expertise, the extended virtue model preserves the objectivist core of the substantialist view but extends and reformulates it in non-individualist terms. How can we conceive competence in a non-individualist sense? Is epistemic competence

3) Interestingly Goldman, who is the most prominent advocate of what we call the substantialist view in the philosophical literature, partly revised his account recently in a functionalist direction by mentioning the expert's "capacity to help others…by imparting to the layperson (or other client) his/her distinctive knowledge or skills" (Goldman, 2014).

PsychOpen GOLD

not, after all, a psychological disposition? The extended virtue model defines competence as a (potentially) extended or socially distributed quality in line with the theses of *extended cognition* (see, e.g., Clark, 1997; Clark & Chalmers, 1998; also Hurley, 1998; Menary, 2006; Rowlands, 1999, 2009; Wilson, 2004) and *distributed cognition* (see, e.g., Clark, 2017; Hutchins, 1995).[4] Socially extended or distributed cognition describe supra-individual cognitive systems which are composed through the ongoing and reciprocal interaction between multiple cognitive agents and/or technological instruments that individually execute the sub-tasks of a unified cognitive process. An extended competence is accordingly the emergent capacity of a supra-individual cognitive system to execute a complex cognitive task. The extended virtue model of expertise can equally apply to individual and socially distributed competences in science, because the concept of a reliable informant does not make any commitment to individuals: *the right informant is the seat of the relevant competences*. Thus, it can incorporate socially extended expertise and genuinely collective experts (such as big research teams).

Being able to properly fulfill the function of an expert also has an equally central *responsibility* aspect. Epistemic responsibility (e.g., Code, 1984; Montmarquet, 1987; Zagzebski, 1996) has to do with a desire to achieve the proper ends of scientific inquiry and scientific communication; particularly, the advancement of scientific knowledge and its credible transmission. Responsible scientific practice is first and foremost *honest*: Making scientific claims with a view to misinform or create a false impression (e.g., scientific fraud or falsification) directly contradicts the function of expertise and obliterates its basis. In the second place, responsible scientific practice is *diligent* and shows vigilance towards sources of error and bias. Negligence of counterevidence, insensitivity towards error, or self-serving biases undermine the credibility of scientific inquiry and thus the reliability of the expert as an informant. Ideally, the expert judgment also manifests other intellectual virtues such as impartiality, intellectual humility, or openness to criticism. Different intellectual virtues can be seen as particular aspects of a certain unitary attitude towards research that is guided by a concern with learning about the world, and lacking any of these to a substantial degree undermines proper conduct of inquiry. To illustrate, a responsible scientist treats alternative perspectives fairly on the basis of the available evidence, is open to the ideas and criticism of other scientists, is willing to conceive and examine alternatives to their own or popular theories, has a sense of own fallibility and is willing to admit mistakes, and does not oversell or miscommunicate their results. This kind of *responsibility* is an essential part of the very content of expertise, because it underlies the very function of expertise, alongside *competence*.

4) In this context we can speak of both object-level competences for producing the relevant kind of empirical evidence for a particular scientific claim, and meta-competences for evaluating the reliability of the outputs produced by object-level competencies. Making sufficiently well-warranted scientific claims requires both.

PsychOpen GOLD

Lastly, epistemic competence and responsibility are closely intertwined in the function of a scientific expert, namely being a reliable informant, because we expect scientific experts to acquire, exercise, and maintain their scientific competences with a view to reliably and credibly advance scientific knowledge. Experts are thus responsible for cultivating and using their scientific competences adequately and honestly when informing others in their domain of expertise. In this regard, a scientist who makes scientific claims that they are not sufficiently competent to research, or one who does not continue to develop their knowledge and skills after achieving a position of authority shows a lack of responsibility, even if they are following the current norms in their fields.

What is Different in the Extended Virtue Model?

The extended virtue account radically diverges from the substantialist view in two respects. First, competence (e.g., relevant special knowledge and skill) is necessary but not sufficient for being an expert, because it is competent *performance* that defines the social epistemic function of expertise and this equally comprises a responsibility/virtue element. Second, neither competence nor intellectual virtues are necessarily individual (for collective intellectual virtues and vices, see, e.g., de Ridder, 2022; Fricker, 2012, 2020; Uygun Tunç & Pritchard, 2022). We can ascribe qualities like responsibility, integrity, diligence or negligence, openness, or dogmatism to groups such as research teams.[5] For instance, a research team can be deemed diligent to the extent that it implements rigorous cross-checks on statistical analyses, team members are vigilant towards each other's potentially problematic use of researcher degrees of freedom, and detected errors are honestly dealt with or admitted rather than concealed. On the other hand, a research team would be irresponsible as a collective body if one team member discreetly fabricates data and no one else shows serious concern about the actual origins of the data.

The extended virtue account diverges equally radically from the implicitist and the social conventionalist views. The implicitist view of expertise does not feature any intersubjectively accessible basis for identifying and evaluating expertise, whether the criterion is uniquely personal implicit knowledge or knowledge implicit to a community of practice, unlike the notion of competent performance. Relatedly, although the implicitist view can be said to maintain competence as an aspect of expertise, it is difficult to speak of epistemic responsibility when the competence in question is implicit. Therefore, unlike the extended virtue account, experts, as defined by implicitist views, cannot be easily subjected to public criticism or held accountable. The social conventionalist view similarly lacks any objective basis, since its criterion for expertise is none other than rec-

5) In the case of research teams, intellectual virtue or vice is a matter of shared attitudes, dispositions, values as well as efficiency in implementing second-order processes of internal scientific criticism, quality control, and error detection.

PsychOpen GOLD

ognition by a large enough social group. This kind of recognition is not always justified, and we readily see in the case of widely recognized "experts" in pseudo sciences.

Astrological "Experts"?

We can illustrate the importance of focusing on informant reliability instead of any other criteria for expertise in isolation through a radical analogy. Let us take expert claims in pseudoscience like astrology. Astrology "experts" claim to foresee future events by closely observing constellations. They may (and probably many do) possess a wealth of knowledge about astrological theories and how these are used to make empirical predictions, as well as significant skill in observing, recording, and interpreting the patterns and movements of celestial objects in the light of astrological theories. The falsity of their predictions does not stem from ignorance about the theories and methods of astrology, or a lack of time-honed talents and skills, or initiation to the defining practices of an intellectual community, or a lack of recognition and esteem from their peers. Thus, none of the traditional perspectives allows us to effectively pinpoint the problem with astrological expertise, unlike the extended virtue account we propose. Astrological experts fail dramatically in fulfilling the function of an expert in the way they claim expert status: They are clearly not reliable informants regarding future events. However, they might be seen as experts on the common theories and practices of astrology. Their expertise may have historical or anthropological value, although it does not have any intellectual value regarding forecasting future events. Where does the value of psychological expertise come from? The social epistemic function of a psychology expert is to produce reliable and credible information about the nature of psychological phenomena for other scientists and the public, predict the future states of such phenomena, develop effective psychological interventions or advise social policy. In cases when experts in psychology fail to fulfill any of these, unfortunately the same conclusions as those about astrological experts would also apply to them.

So, in contrast to the more traditional perspectives on expertise, the extended virtue account can pinpoint what is problematic with expert claims in astrology, and still more usefully, allow us to identify how our definition of expertise facilitates or hinders efforts to increase the capacity of psychological science to self-correct. We argue that the prevalent endorsement of criteria for expertise that derive from some older theoretical perspectives currently hampers the reform movement by leading to resistance to reform-minded ideas on topics such as the need for and the value of direct replications.

How to Evaluate Psychological Experts as Reliable Informants?

We said that psychological experts can be considered reliable informants to the extent that they provide reliable and credible information about psychological phenomena, make accurate predictions about these phenomena, develop and implement effective interventions for psychological problems, or contribute to developing more effective

PsychOpen GOLD

social policies for societal issues. How best to measure success in these areas is beyond the scope of this paper. But there are two important points that we think should not be overlooked when formulating such measures.

First, we have to note again that experts do not have to be individuals. Depending on the scientific topic, acting as a reliable informant may require a single person, a research consortium, or even a whole group of people working in a particular discipline. So, the assessment of expertise on a topic should not necessarily be based on metrics developed to assess single individuals. We can think about replicability rates and the provision of reliable and credible information about psychological phenomena in this context. To say that a field with a very low replication rate can provide reliable and credible knowledge about its subject matter is hardly possible, and the replicability of individual studies requires competent and responsible behavior on the part of the person or persons carrying out the studies. But in many cases, individual competence and responsibility may not be enough to increase (or even test) replicability. In these cases, it is necessary to work with a group of individuals possibly in a consortium or an adversarial collaboration. In addition, if a certain number of replication studies are not conducted in a discipline, it is not possible to know to what extent the field produces reliable information. For these reasons, when evaluating any claim of expertise, depending on the nature of the subject matter, any of the replicability indices of the researchers themselves, the research group they work with, or the sub-discipline and the discipline in which they are involved may come to the fore.

The second issue we want to briefly point at is that metrics should be directly geared toward capturing a track record of competent and responsible epistemic performance. The currently popular metrics (e.g., *h* factors) for assessing expertise can only be very indirect and unreliable indicators of epistemic performance, as they are meant to measure expertise in a social conventionalist sense. As steps in the right direction for targeting epistemic performance, we can give the curation framework (LeBel et al., 2018), false discovery risk indices (Schimmack, 2022), or empirical replicability audits (O'Donnell et al., 2021) as examples.

In the following, we further elaborate on the main points of contrast between the traditional perspectives on expertise and the extended virtue model by discussing how the mainstream or established understanding of expertise in psychological science (which bears the influence of these traditional perspectives) drags the reform movement down. In particular, we will discuss how some crucial issues that are on the reform movement's agenda such as big team science, perverse incentives, and fraud/questionable research practices are all associated with how we define and model expertise.

PsychOpen GOLD

Expertise and the Problems of Self-Correction in Psychology

Fraud and Questionable Research Practices

The credibility of science shapes (at least to a significant degree) public trust in science. Still more crucially, some degree of trust is in the fabric of contemporary scientific knowledge production.[6] Scientists can produce knowledge only interdependently; namely, by sharing their findings and building on the research of others, through specialization and division of cognitive labor, or by closely cooperating on the same projects. For this reason, knowledge or skill asymmetries are ubiquitous in science and most scientists are in an expert-lay relationship with scientists in other fields or even sub-fields. At this level of interdependence, the success of the whole scientific enterprise hangs on the trustworthiness of scientists as information providers, i.e., scientific experts. A scientific expert is trustworthy to the extent that they are *reliable* as *informants*; that is, to the extent that the others who make use of their research are informed accurately sufficiently often (or misinformed sufficiently rarely).[7] A reliable epistemic performance in this context can comprise successfully predicting events, offering credible and informative explanations of phenomena, identifying falsely accepted scientific claims, or establishing novel facts. If scientific experts are not reliable enough, the role trust plays in science would only be that of undermining the credibility of science and ultimately eroding the basis of public trust in science.

However, whether we ascribe expert status on the basis of possessed knowledge and skills (explicit or implicit), faithful reproduction of received research practices, or eminence and reputation, we divorce the concept of expertise from its central *function* of acting as a reliable informant in a scientific domain. Negligence of the very function of an expert, and consequently, a lack of attention to how reliably one performs as an expert is a characteristic weakness of all traditional perspectives on expertise (i.e., substantialism, implicitism and social conventionalism). It is thus worthwhile to start with a case that illustrates the failure to perform the function of an expert, and hence indicates a central problem with all these traditional perspectives; namely fraud and questionable research practices.

6) Compare de Ridder (2022), who maintains that scientists do not trust other scientists to be reliable informants (which is considered too strong), but merely to stick to the prevailing standards of scientific practice. If the prevailing standards are sufficiently high, scientists would in effect be reliable informants. However, if they are too low, there is no point in "trusting" other scientists to stick to them—this can only be a reason to distrust most other scientists. That "trust among scientists is inevitable", as de Ridder rightly assumes, means essentially that scientists must build on one another's findings and ideas, and they do so inevitably on the basis of some degree of trust. If scientists do not trust their peers to be reliable informants, there is no rational point in building on their research and putting own credibility also at risk.

7) This includes enabling others to rightly suspend judgment when there is insufficient or mixed evidence.

PsychOpen GOLD

In the substantialist and implicitist views of expertise, one's expert status has nothing to do with acting as a reliable informant, which we put at the center of the extended virtue model. Since the basis for expertise ascription in these perspectives is only one's relevant knowledge and skill, they do not feature any epistemic *responsibility* criterion that is intrinsic to their definition of expertise. Expert status is gained only by virtue of acquiring the relevant knowledge and skills, i.e., competences. However, an expert´s knowledgeability and skill do not directly imply that their scientific claims will be reliable, since these qualities are compatible with lack of honesty and transparency such as fraud or selective reporting. Thus, scientists' knowledge and skills should be seen as constituting expertise only to the extent that these are properly manifested in the reliability and rigor of the research procedures they follow and in their ability to critically analyze and interpret research results in their field, so only if they show a high-quality epistemic performance as informants. The definition of expertise from a social conventionalist perspective similarly lacks any reference to responsibility (while also lacking any reference to competence). Eminence, reputation or popularity are in principle all compatible with engaging in questionable research practices or even fraud.

Since all these perspectives ignore epistemic responsibility as an informant, the received view tends to treat matters such as scientific integrity and research ethics as conceptually external to scientific expertise. These are regarded to have little to do with a researcher's cognitive/epistemic qualifications to claim expert status. The valuation of eminence or reputation (deriving from the social conventionalist view) functions in practice as yet another barrier to accountability. Unless severe misconduct is admitted/demonstrated (although nearly impossible), the expert status of an eminent scientist is never undermined. Even when that happens, we blame their morals not their core qualifications and thus expert status: We regard them as immoral experts but as experts nonetheless.

An illustrative example of this is one of the incidents that triggered the credibility crisis in psychology. Diederik Stapel was both a deeply knowledgeable and skillful social psychologist and a highly eminent academic figure who held numerous distinctions and awards. In 2011 it became known that he committed fraud in dozens of published papers.. Stapel competently formulated hypotheses and predictions, explicated their theoretical background, designed experiments which he did not actually run, and analyzed data he never collected (Levelt Committee, 2011). Should we consider Stapel an expert? In the social conventionalist view, he was clearly an expert even while he fabricated data and would remain one if he did not admit to it. Eminence creates a significant degree of protection from criticism by increasing trust and reducing skepticism, thus making it more difficult to identify fraud. In the substantialist view (and its implicitist variant), again he would most certainly be considered so, because he has extensive knowledge and a high level of skill regardless of the actual origins of his data. Stapel retained his extensive knowledge and experimental skill despite consistently engaging in misconduct,

PsychOpen GOLD

and still does. Stapel did not fail to cultivate excellent skills or build a reputation; he failed only in properly fulfilling the function of an expert: informing reliably in the domain of expertise. Stapel put this knowledge and skill into *misinforming* others instead. His results were much more misleading than random error due to his skill in data fabrication, which ensured that they did not raise suspicion.. Most importantly, as an epistemic authority who is supposed to be a reliable informant, he actually diminished others' chances of finding true answers to those questions by themselves.

The maintenance of competence or eminent credentials has thus nothing to do with having a track record of reliable and credible expert reports. Neither misconduct nor engaging consistently in weak or unreliable research practices undermine one's core scientific competence or credentials. Thus, if we stick to the traditional views of expertise, once gained, the expert status is very difficult to lose. None of these views gives us substantial, theoretically justified reasons to disqualify Stapel as an expert, apart from the contingent fact that he actually admitted fraud. We should also bear in mind that revealed fraud cases like Stapel's are extremely rare.

In comparison, according to the extended virtue model one does not become an expert indefinitely, but only for as long as and to the extent that one properly fulfills the function of an expert. When we conceive expertise as informant reliability, then the impropriety of making dishonest, incomplete or potentially misleading scientific claims is suggested by the very concept of expertise. In other words, a scientist whose reports are too unreliable too often simply should not be regarded as an expert. Scientists gain and maintain an expert status only through building and maintaining a track record of successful and reliable performance as information providers in a scientific domain. This implies that their scientific outputs would be open to consistent scrutiny for scientific quality and expert status would never imply protection from criticism.

Collaborative Research and Collective Expertise (vs. Substantialism)

An individualist conception of the scientific expert is becoming less and less viable with the increasingly collective outlook of scientific knowledge production. The collective character of science is particularly prominent in large interdisciplinary research collaborations which we see today at the frontiers of science, from biology (Rosenstein et al., 2014) to physics (Aad et al., 2008). Research collaborations not only have at their disposal a much wider scope of domain expertise to effectively tackle interdisciplinary questions, they are also often in a better position to minimize sources of error than individual researchers, because they can better check for possible confounding factors, counter the experimenter's bias or improve sample selection[8]. Besides their methodological benefits, team science allows for plurality in theoretical and methodological perspectives to be applied in tackling complex questions (Tebes et al., 2014). There have also been recent calls for big team science in psychology (Forscher et al., 2020; Uhlmann et al., 2019). We

PsychOpen GOLD

also have already seen a few successful collaborative projects under the name Many Labs (Ebersole et al., 2016; Klein et al., 2018, 2022), Many Babies (Byers-Heinlein et al., 2020), Many Smiles (Coles et al., 2022), and also the initiative Psychological Science Accelerator (PSA), which is a crowdsourced research network consisting of 1328 members in more than 84 countries (see also Moshontz et al., 2018).

But, if research collaborations are so well-positioned to tackle a variety of theoretical and methodological challenges, why do we not see them more often? Especially since there seems to be a felt need for larger research collaborations in psychology to investigate the complex questions that human behavior poses, why do the calls for big team science not find a wider appeal, and why do the existing initiatives face significant practical challenges? Part of the reason is the conceptual challenge of imagining a collective expert: How to identify, evaluate, credit, reward or hold accountable an expert who is not an individual?

A large interdisciplinary research collaboration implements a complex research plan that requires diverse expertise, simultaneous manipulation of multiple scientific instruments, or data collection at different times and places. Implementing such complex procedures exceeds the cognitive ability and capacity of individual researchers and requires a larger cognitive system comprising multiple agents. The individual scientist is only a contributor to the production of this kind of information, because they will at best have a superficial understanding of the total evidence (including higher-order evidence of reliability, validity, generalizability, etc.). The only entity that has the relevant expertise for the production and assessment of the total scientific evidence is thus *the research collaboration as a whole*. For this reason, it is more appropriate to conceive of large research collaborations as a single (supra-individual) informant, to identify them as "the expert." A most concrete sign of this is collective-authorship by consortia (see, e.g., Fontanarosa et al., 2017), which implies substantial changes in how credit and responsibility are allocated in science—hence in how we ascribe expert status.

However, it is conceptually very difficult to think of collective experts from a substantialist view, because its criteria for expertise are defined in mentalistic terms. Knowledge and skills are qualities of individuals, because they are strictly psychological. A notion such as a group expert is ill-conceived, because it seems to stipulate a supra-individual psychological entity like a group mind. The substantialist view of expertise relies on a traditionally individualistic conceptual foundation in epistemology, according to which knowledge as well as competence to produce knowledge are states that are intrinsic to an agent's mind. This kind of individualist epistemology faced a substantial challenge in the last decades by the theories of extended and distributed cognition, which make a case for the existence of veritable supra-individual cognitive systems. While of

8) Cf. Winsberg et al. (2014), who are more skeptical about accountability in big research teams. For a defense of the strengths of research collaborations in this regard, see Uygun Tunç (2023).

PsychOpen GOLD

some of these also argue for (technologically and/or socially) extended minds (see, e.g., Clark & Chalmers, 1998), a more modest thesis that finds widespread appeal is that there are veritably supra-individual cognitive competences and cognitive performances (e.g., de Ridder, 2014; Giere & Moffatt, 2003; Palermos, 2022; also Uygun Tunç, 2023). For our purposes, we can meaningfully speak of group competences and group responsibility, and this suffices to speak about group experts.[9] What matters the most in the context of expertise from the extended virtue perspective we propose is that collective entities can in some cases function as more reliable informants than individuals ever can. Research collaborations are a prime example of this for the reasons we briefly explore in the following, and psychology in particular can greatly benefit from their prevalent existence.

In the context of psychological science, research collaborations seem to be particularly suitable for meeting difficult methodological challenges associated with producing higher-order scientific evidence such as replication studies and studies that investigate the generalizability of an existing finding (see Forscher et al., 2020). Multi-lab, multi-site replication studies indeed efficiently serve to increase statistical power, and check for the possible confounding effects of experimental settings, population characteristics, researcher bias, or sample selection (Moshontz et al., 2018). One can argue that wider-scale collaborative projects are almost a necessity in order to increase the generalizability, validity, rigor, and reliability of psychological findings (Munafò et al., 2017; Nosek et al., 2022).

To illustrate through a discipline-specific challenge to generalizability, we can think of the fact that cultural influences have a substantial impact on human behavior (see Lehman et al., 2004). Studies show that an overwhelming majority of published psychology studies are conducted with US-based samples by US-based authors (Arnett, 2008; Thalmayer et al., 2021). It goes without saying that if psychology wants to be the universal science of human behavior, it must expand its cultural horizons outside of the US (see Adetula et al., 2022; Forscher et al., 2021; IJzerman et al., 2021; Silan et al., 2021). To achieve this, psychology needs big research teams involving researchers from diverse cultural backgrounds.[10]

There is an undeniable need in psychology for big team science in producing first-order evidence too. The reform literature is rife with studies that show many published papers in psychology to have serious defects in the way they a) define their theories, concepts, and hypotheses (Bringmann et al., 2022; Meehl, 1990; Oude Maatman, 2021),

9) For a more detailed discussion, see Uygun Tunç (2022); for a discussion of why group competence does not entail group mind, see Uygun Tunç (2023).

10) We can see the same problem only reproduced in the example of neo-colonial (or helicopter) science (Dahdouh-Guebas et al., 2003). Due to the tendency to think of expertise in substantialist terms, we expect a researcher with a WEIRD (Henrich et al., 2010) background to effectively formulate a theory, design a severe test, and correctly interpret the results in isolation (or in small teams at best) regarding a behavioral phenomenon that takes place in a completely different culture. In many cases, this results in false analogies and overgeneralizations.

PsychOpen GOLD

b) make their measurements (e.g., Flake & Fried, 2020; Weidman et al., 2017), c) plan and implement correct statistical models (e.g., Bakker et al., 2012; Crede & Harms, 2019; Rohrer et al., 2022). There have been multiple reasons suggested as to why this is the case (see Nosek et al., 2022), and most probably each explains an important portion of the problem. Another probable reason that we put forward here is that psychological experts as isolated individuals (conceived in line with the substantialist view) simply lack the necessary competences to be reliable informants in their domain of "expertise." This is because their domain of expertise requires a very diverse set of competences that can only be found in bigger research teams at satisfactory levels. Due to its individual-centered organization, psychology still requires its experts to be polymaths who need to have profound knowledge in a very diverse domain of subjects alongside their immediate subject matter. The reformist criticisms of the current practices in psychological science might also be indicative of the fact that, for some questions at least, this is an unrealistic expectation.

Lastly, team science can also be a uniquely effective way to produce high-quality first-order evidence on contested research topics. In this context, we can mention another, more special case of scientific collaboration which is directly applicable to psychological science: adversarial collaborations (Mellers et al., 2001; Tetlock, 2006). The scientific literature on highly contested research questions typically contains substantial empirical support for several opposing claims. Such contested research questions may be extremely difficult to definitively answer, because often the parties endorse incompatible theoretical assumptions, lack common criteria for falsifying points of view, and disagree on key methodological issues (see also Uygun Tunç & Tunç, 2023). Adversarial collaborations are a very suitable solution to deal with these kinds of situations.

Low Replicability and Scientific Rituals (vs. Implicitism)

Replicability vs. Implicit Expertise

When Baumeister (2016) said that "the optimal approach for now is not a matter of replacing an obsolete, ineffective system with a shiny new one. Rather, we should continue business as usual while adding new approaches… keep the old model alongside the new model(s), rather than replacing it," he was not only referring to the established practices of research, publication, funding and promotion, but also (maybe more implicitly) to the traditional understanding of a successful scientific expert in social psychology. Namely, someone who has "an intuitive flair for how to set up the most conducive situation and produce a highly impactful procedure," which is acquired by "years of specialized training and skill cultivation." He contrasted the "flair" of a good experimentalist to the graduate students in his lab who fail to replicate his studies due to incompetence - the "bad experimentalists" who must seek other careers.

This conception of a good experimentalist derives its theoretical justification from Polanyi's notion of implicit or tacit knowledge. Polanyi (1966) argued that "the paradig-

PsychOpen GOLD

matic case of scientific knowledge… is the knowledge of approaching discovery. To hold such knowledge is an act deeply committed to the conviction that there is something there to be discovered. It is personal, in the sense of involving the personality of him who holds it, and also in the sense of being, as a rule, solitary." We see here a profoundly different conception of scientific activity than a focus on rigorous and critical investigations of the empirical merit of theories—as in Popper (2002a, 2002b), for instance. Here the focus lies on the unique personalities of accomplished scientists, their unforeseeable and unexplainable bouts of intuition or "Eureka" moments as the main locomotive of science. This perspective is explicitly positioned against (primarily, by Polanyi himself) a family of notions that characterize scientific objectivity such as intersubjective testability, explainability, formalizability, and replicability. If scientific knowledge is the product of creative and inexplicable, unformalizable processes which are shaped by uniquely personal qualities, then demanding replicability is both unreasonable and unrealistic. Underlining this connection, Derksen (2017) maintains in a similar vein that "you could never fully spell out how to replicate an experiment," because some experimental skills are "implicit – not written down in textbooks or in the method section of reports, not explicitly taught."

It is indeed true that most scientific studies require expertise in a particular method or domain of research to be properly conducted. That being said, intersubjective testability or replicability is one of the core scientific values in confirmatory research, by virtue of being the most minimal criterion a scientific practice must satisfy in order to qualify as reliable. Endorsing replicability as a core value means that the term "skill" can be understood only as objectively defined adequate experience and training, which might amount to having an academic qualification (e.g., a degree in a relevant field) or demonstrable expertise in certain techniques (e.g., demonstrable previous experience with the online experiment software). Unless one does not forgo the claim for intersubjective testability and thus objectivity altogether in the context of confirmatory research, as well as associated demands for scientific integrity and accountability, implicit knowledge and skills cannot be part of the definition of expertise.

Intersubjective testability or replicability is also closely tied to the notion of scientific criticism and the critical rationalist approach to scientific self-correction (see Popper, 2002a, 2002b). The expertise claim of the individual scientist from the implicitist perspective can be reliable only to the extent that they are able to critically evaluate their own work and all the relevant previous work in the field. This is not only unrealistic but also overlooks the fact that nobody, however virtuous, is the best critic of their own work due to cognitive and theoretical "blind-spots" (see also Longino, 1990). This challenge to scientific objectivity or impartiality is much better met by a critical culture where everybody criticizes the work of others. The cornerstone of a critical culture is that scientists manifest openness to criticism. Implicitism about expertise, however, means by definition that some aspects of scientific inquiry cannot be open to scrutiny, even

PsychOpen GOLD

by the researchers themselves. Further, community criticism is not a viable notion if we do not assume that all reliable methods and practices in confirmatory research are in principle repeatable. Consequently, we also cannot rationally argue for a responsibility or duty on the part of researchers to either publish replicable studies or to check whether published studies are replicable. Thus, scientific self-correction through mutual criticism is categorically incompatible with implicitism, while implicitism itself does not have any viable strategy for promoting reliability.

Scientific Rituals and Communities of Expertise

Scientific skills may also be seen as implicit within communities of scientific practice rather than individuals, which is equally problematic. Collins and Evans (2008) argue for a version of implicitism that focuses less on unique personal qualities and more on communities of practice. They maintain that an expert in a scientific field is someone who can (at least) meaningfully interact with others in that field ("interactional expertise") or contribute to that field ("contributory expertise"), in the sense of a competence to engage in the kind of characteristic (potentially esoteric) activities they do.

The blind spot of this kind of view is that scientific practices may be widely accepted as standard in a field but can nonetheless seriously lack reliability (i.e., lacking robustness, validity and so forth). The implicitist view of expertise lacks any criterion pertaining to the actual reliability of an expert's performance as an informant. In this respect it is vulnerable against a common criticism by the scientific reform movement in psychology of "ritualized" research practices (see Gigerenzer, 2004). These are the kind of choices researchers make in the course of a study that are justified not on the basis of theoretical, empirical or philosophical reasons but merely because the relevant choices are the social norm in that field. Ritualized research practices can range from the relatively trivial "How to phrase a hypothesis?" to the more substantial "How best to experimentally manipulate emotions?" or "Which potential confounds to consider in measuring attitudes?"—the more the answers to these questions lean in the direction of "This is how everyone does it", the more we are speaking of a scientific ritual.

The implicitist account of expertise cannot objectively distinguish between an actually reliable research practice and a merely ritualized one. Following the criteria proposed by Collins and Evans (2008), we can say that a researcher can possibly count as manifesting interactional and even contributory expertise by merely mimicking the execution of some protocols, with little to no critical reflection.

Perverse Incentives (vs. Social Conventionalism)

Perverse incentives are arguably one of the most popular discourses in the reform literature (e.g., Chambers, 2017; Munafò et al., 2017; Nosek et al., 2012; Simmons et al., 2011; Smaldino & McElreath, 2016). In the context of scientific expertise, the adjective 'perverse' might imply that incentives do not track real expertise. The current incentive

PsychOpen GOLD

structures in science are in fact perfectly aligned with a social conventionalist understanding of expertise, according to which somebody is a scientific expert only if and to the extent that they are recognized as one by a reasonably big portion of the scientific community, in accordance with their standards. Thus, the best indicator of expertise in this sense is eminence or reputation. From this perspective, it is meaningless to question if eminence actually indicates real expertise—if they do not, by whose standards? This kind of question is arguably at the bottom of the reaction from the established research tradition in psychology to the talk of perverse incentives.

Incentives do indeed play a big role in shaping and maintaining certain scientific practices. We believe that the crisis of self-correction may be less due to the inability of the current incentives to track expertise per se, but rather to the type of expertise that is being incentivized. At present, expert ascriptions are often made on the basis of *credentials* such as education, affiliations, publications at respected venues, collection of awards, and grants. The way the current incentives are set does not consider the collective self-correction of science as a central feature of scientific expertise. That is, the incentivization of eminent credentials aligns with a goal of maximizing groundbreaking individual research projects, which unavoidably neglects the necessarily collective nature of scientific self-correction (see also Vazire, 2017). The metrics that we currently use in the assessment and incentivization of expertise are almost completely shaped by a focus on eminent credentials, as can also be seen in their historical development (see Bourdieu, 2004). So, when the current incentives fail to facilitate self-correction and a more credible psychological science, it is not because they are somehow corrupted after some time. The current incentives fail to facilitate self-correction because they are not designed with that aim in mind. The current incentives are not "perverse"—they do what they are supposed to do in accordance with a social conventionalist conception of expertise (albeit, sometimes imperfectly); that is, they incentivize eminence rather than informant reliability.

Even in a scientific utopia where we find ungameifiable metrics to base our incentives on, we will be incentivizing individual eminence in one way or another as long as we adhere to a social conventionalist conception of expertise. The problem with this is again mainly twofold: 1) the neglect of the social epistemic function of expertise, which consists in serving as reliable informants, and (as a consequence of this neglect), 2) the exclusive focus on individuals rather than a social organization that would optimally fit the scientific tasks and challenges at hand (self-correction in psychology, in particular, requires a collective effort, as was perfectly exampled by Many Labs studies).

Seeking scientific knowledge is an ongoing effort by a whole community of inquirers. Parsing this distributed, often complicated, and messy process at novel and surprising findings gives us a misleading picture of scientific activity that neatly foregrounds individuals and their achievements. The current incentives which favor eminent credentials indeed fit quite smoothly with this picture of science. In terms of academic career

PsychOpen GOLD

advancement, the most important incentives are the ones related to publishing, and these incentives are known to prioritize novelty and surprise over epistemic values that are more closely associated with truth, such as accuracy or generalizability (Nosek et al., 2012). There is a selective advantage in terms of publishability for the papers that tell a simple, novel, and surprising story (Giner-Sorolla, 2012). Studies that tell a more complicated story because of their more diligent methods are less cited and harder to publish, so incentives motivate people to be strategically vague in designing/implementing/reporting a study (see Frankenhuis et al., 2022).

Close replication studies, which are essential for self-correction (see Simons, 2014; Zwaan et al., 2018), are also relatively rare (Hardwicke et al., 2022; Makel et al., 2012) and share the same fate with studies that present more complex but more reliable and credible findings resulting from diligent methodological procedures. The scarcity of close replication studies is closely related to an understanding of scientific expertise that ignores its social function, namely serving as a reliable informant. A successful replication is actually a piece of valuable meta-information on informant reliability. However, due to the focus on eminence over being reliable informants, close replication studies, which cannot contribute to eminence, are not seen as a necessary part of the knowledge building process.[11] Thus, incentives do not typically target those features of science which do not make the best story but are essential to collective, ongoing inquiry after scientific truths, such as checking for replicability, statistical reproducibility, robustness against various sources of heterogeneity, or devising increasingly more severe tests for the contending theories in a discipline. These are all part of a culture of ongoing criticism, also called critical rationalism by Karl Popper (2002a, 2002b).

The social conventionalist view of expertise with its emphasis on eminent credentials either neglects the demands of or directly contradicts a more collectivist and critical rationalist research culture, which largely characterizes the aspirations of the reform movement in psychology. This is because when we rely on eminent credentials, we unavoidably grant individuals a position of intellectual authority that might not always be warranted. This intellectual authority works as a form of justification itself, so we require less independent scrutiny to trust the scientific claims. For example, some recent studies on peer review practices provide empirical support that reviewers are indeed more lenient towards more eminent researchers (Tomkins et al., 2017). Under this culture of epistemic authority based on eminence, it is not surprising that some authors feel threatened by close replication studies conducted by independent teams. Their expertise claims do not depend on replicability or other indicators of informant reliability. Then, for example, close replications are at best a nuisance, or worse: a counterclaim to the

11) Some of the resistance to close replications are due to the fact that some scientists adhere to the approaches in philosophy of science that puts less emphasis on close replications. Yet, we believe that such principled opposition to close replications in psychology is comparatively rare.

PsychOpen GOLD

author's status as an expert. Because, why do you need to replicate these findings? Do you not find their credentials persuasive enough? It is not surprising then that the scientific reform movement in response to the credibility crisis has experienced normative clashes. In a research culture organized around eminent individuals, it is counter-normative to conduct independent replications without the cooperation of the original author or to go through the methods and results in the published literature and call out a serious suspicion of QRPs or data fraud. People who did so have even been labeled "methodological terrorists."[12]

The rarity of close replications is not just a problem for those researchers who produce unreplicable findings. A group of experts who use a knowledge base that depends on unreplicable observations can hardly be reliable informants. Therefore, insufficient number of close replications also create a collective problem of expertise. An individual-centered understanding of expertise further complicates the problem by failing to provide a clear account regarding who has the responsibility for conducting replications. In any individualist model of expertise, the expert should be a relatively self-contained unit of knowledge production. A need for close replications conducted by independent labs, on the one hand, is something that directly contradicts such a model, as it indicates strong interdependence. Achieving a high proportion of replicable observations in the literature can only be a collective task that is distributed to all the researchers working on the relevant topics. Since the task is collective, responsibility must also be modeled collectively. In Islamic jurisprudence a distinction is made between individual duties (fard al ayn) that must be individually fulfilled by all believers and community duties (fard al kifayeh) which, when performed by only some of the believers, free all believers from responsibility. For example, daily prayer is required to be performed by everyone, though funeral prayer is only required to be performed by a sufficient number of people in the community and not everyone is expected to participate in it. Analogically, replication studies should also be seen as community duties. For an effective process of self-correction to work, it is not necessary that each individual study is replicated or that each individual scientist regularly conducts replications. If a large enough number of replication studies are being done, the expertise claims of all the researchers are supported by it regardless of whether they conduct such studies or not. But if a large enough number of replication studies are not being done in a field, then this undermines the claims of expertise for every researcher in the field.

The idea of community duties or any other model of collective responsibility can only be an oddity in an individualist model of expertise. However, a coherent account of collective responsibility is a must for replication studies and thus for effective self-correction in science. The extended virtue model not only places equal emphasis on epistemic

12) See https://www.thecut.com/2016/10/inside-psychologys-methodological-terrorism-debate.html

PsychOpen GOLD

responsibility as on competence, but also allows for a socially extended or distributed conception of expert responsibility.

In the case of collective responsibility, the risk is a diminished sense of individual duty due to the division of responsibility. This risk can be mitigated by concretely defining the duties of all relevant subjects, from researchers to scientific journals and research institutions. If conducting replication studies is modeled as a collective responsibility, one way to do this might be to require each researcher to maintain a certain ratio of original to replication studies. Journals may be required to reserve a publication quota for replication studies, or dedicated journals may be established for replication studies. On the publication end, meta-analyses could incorporate information on how many independent replication studies support (or undermine) each finding. A more substantial intervention may be to build a number of systematically linked replication studies into the original test of hypotheses (see Uygun Tunç & Tunç, 2023), at least when testing contested research questions or where there are significant doubts regarding the validity of the methods and the generalizability of the findings.

We cannot stress this enough: A definition of expertise that is based on eminent credentials rather than epistemic performance inescapably creates an incentive structure that contradicts the goal of self-correction in science. We need to have a more performative understanding of expertise to incentivize more effective self-correction in science.[13] Therefore, we must change how we *understand* expertise, not merely how we *measure* and *incentivize* it.

Discussion

Sticking to the criteria for expertise promoted by one or more of the traditional perspectives we discussed is arguably not bereft of any rationale. To paint a more balanced picture, first of all, an individualistic conception of expertise (which characterizes most of these traditional perspectives) is more in line with traditional epistemology and the institutional organization of science. Individualism still characterizes mainstream epistemology, thus we think of knowledge, competence, epistemic responsibility, and intellectual virtues as pertaining categorically to individuals. In accordance with this, the social organization and the institutional values of science, science policy as well as

13) It might be questioned if reliable performance is indeed an objective criterion, and not socially determined just as eminent credentials. It might be argued, after all, that evaluating reliability would require certain standards, and standards are social conventions. While there is some truth to this kind of objection, nature does not tell us directly which scientific claims are true, but instead the scientific community collectively decides to accept or reject them as facts. But acceptance is always tentative, and the scientific community, to the extent that it upholds an empiricist and critical culture, can always change or revise a consensus in the face of novel observations. Rejecting this about science is tantamount to rejecting its status as a special institution, which one does only if they endorse a radically social constructivist view of scientific knowledge.

PsychOpen GOLD

the scientific credit/reward economy are designed around individual experts. While one of the most exciting and promising developments in contemporary epistemology is the proliferation of non-individualist accounts of all these,[14] there is arguably still time until radical epistemological individualism is a thing of the past and notions such as collective competence or group knowledge are part of the consensus.

Moreover, traditional models of expertise have a stronger affinity with certain values that the scientific institution historically cherishes, such as intellectual autonomy and creativity. Scientists are traditionally given almost complete freedom (and little accountability) and we tend to refrain from evaluating their epistemic performance because we deem this the best way to achieve a highly productive and progressive, consistently ground-breaking science. A belief in genius, unique cognitive gifts, tacit knowledge, or individual flair goes naturally well with this goal, as well as minimal monitoring or central planning.

Moreover, the scientific enterprise itself serves multiple aims. To explicate, scientific inquiry broadly consists of a context of discovery and a context of justification, which reflect the distinction between the generation of hypotheses and exploration of novel phenomena, and their empirical testing and verification (Reichenbach, 1938). These two aspects of scientific inquiry are also associated with two different sets of values, which may sometimes come into conflict (Uygun Tunç & Pritchard, 2022). From the perspective of discovery, there are no objectively superior methods, procedures, or strategies that can be implemented by anyone with a high likelihood of success. The road to discovery is variable, thus distinctive personal qualities such as creativity, perseverance, intellectual courage, or even ambition and recognition-seeking may be positively rewarded. Consequently, scientific success is traditionally believed to be unpredictable and largely due to individual excellence. The opposite is true for scientific justification. In the context of justification, we desire to minimize researcher degrees of freedom, rather than increasing them, because flexibility in the interpretation of results invites bias, noise, and error. Accordingly, we value qualities such as diligence, rigor, vigilance towards error and bias, impartiality, honesty, disinterestedness, and so forth, which are all intellectual virtues associated with epistemic responsibility.

The existence of multiple aims generates the possibility of trade-offs. Safer paths to success (i.e., epistemically responsible and accountable research) are likely to be slower and costlier. Having larger units of scientific inquiry (i.e., individuals vs. research teams) also means fewer attempts at success because there will be fewer hypotheses that can actually be tested. Such trade-offs may not ultimately have an objective solution, due to the involvement of potentially incompatible value systems. Some people may think that

14) The two main pioneers of non-individualist epistemology are the work on epistemic dependence and testimony (see Hardwig, 1985, 1988) and externalist approaches in the philosophy of mind such as the extended and distributed cognition theses, which gave rise to (socially) extended epistemology (Carter et al., 2018a, 2018b).

PsychOpen GOLD

self-correction is a lesser goal in science, because they are more interested in discovering new truths than avoiding errors. When self-correction is not a priority, practices such as performing close replications may seem to have marginal information value. Indeed the culture of science in many fields, especially in psychology, has largely been shaped by a lopsided focus on discovery over justification. On the other hand, we can argue that what makes science a distinctive path to knowledge (among others such as philosophy, religion or art) lies chiefly in the rules and norms governing the practices of empirical confirmation and validation, rather than the processes whereby novel ideas are formed. Thus, it is arguably much less meaningful to speak of a scientific expertise in formulating hypotheses or developing theories, than a scientific expertise in the procedures of empirical justification.

In this regard, scientific self-correction seems to be part and parcel of scientific inquiry and the very notion of scientific expertise. But it is important to note that scientific self-correction is a collective enterprise, so it is a question of how the practices of empirical justification are shaped, validated and monitored by the scientific community. Contemporary science poses novel challenges for systematic self-correction. For this reason, a more social understanding of scientific expertise (with the associated qualities and duties) may be a particularly contemporary need. Until a century ago, when the scientific institution was still operating largely as a gentlemen's club, it was easier for a small number of colleagues to check each other's scientific claims, and implement effective self-correction.[15] Since the scientific community was much smaller back then, it was also the case that substantially fewer hypotheses/theories were tested at any given time. But for today's industrial-scale scientific production, the quality control systems developed for workshop-type scientific production do not work anymore. Thus, from the perspective of systematic self-correction of science, it is not feasible to endorse an individualist notion of expertise and an atomistic understanding of epistemic autonomy. Just as the need for scientific cooperation and interdisciplinary collaboration is becoming more apparent at the frontiers of many fields today, we need an increased emphasis on the epistemic responsibilities of the scientific expert and a more collectivistic understanding of the duty to engage in scientific correction.

Funding: This work was partially funded by the European Union and the Turkish Scientific and Technological Research Council under the Horizon 2020 Marie Skłodowska-Curie Actions Cofund program Co-Circulation2.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

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

15) For instance, in its earlier days a central mission of the Royal Society (and of similar institutions such as the Accademia del Cimento) was to design and conduct replication studies to assess the credibility of reported findings.

PsychOpen GOLD

References

Aad, G., Abat, E., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abi, B. A., Abolins, M., Abramowiz, H., Acerbi, E., Acharya, B. S., Achenbach, R., Ackers, M., Adams, D. L., Adamyan, F., Addy, T. N., Aderholz, M., Adorisio, C., Adragna, P., ... & Batusov, V. (2008). The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation, 3*(8), Article S08003.

Adetula, A., Forscher, P. S., Basnight-Brown, D., Azouaghe, S., & IJzerman, H. (2022). Psychology should generalize from—not just to—Africa. *Nature Reviews Psychology, 1*, 370–371. https://doi.org/10.1038/s44159-022-00070-y}

Agnew, N. M., Ford, K. M., & Hayes, P. J. (1997). Expertise in context: Personally constructed, socially selected and reality-relevant? *International Journal of Expert Systems, 7*(1), 65–88.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602–614. https://doi.org/10.1037/0003-066X.63.7.602

Bakker, M., Van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*(6), 543–554. https://doi.org/10.1177/1745691612459060

Baumeister, R. F. (2016). Charting the future of social psychology on stormy seas: Winners, losers, and recommendations. *Journal of Experimental Social Psychology, 66*, 153–158. https://doi.org/10.1016/j.jesp.2016.02.003

Bourdieu, P. (2004). *Science of science and reflexivity*. University of Chicago Press.

Brewer, S. (1998). Scientific expert testimony and intellectual due process. *The Yale Law Journal, 107*(6), 1535–1681. https://doi.org/10.2307/797336

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, 31*(4), 340–346. https://doi.org/10.1177/09637214221096485

Byers-Heinlein, K., Bergmann, C., Davies, C., Frank, M. C., Hamlin, J. K., Kline, M., Kominsky, J. F., Kosie, J. E., Lew-Williams, C., Liu, L., Mastroberardino, M., Singh, L., Waddell, C. P. G., Zettersten, M., & Soderstrom, M. (2020). Building a collaborative psychological science: Lessons learned from ManyBabies 1. *Canadian Psychology / Psychologie Canadienne, 61*(4), 349–363. https://doi.org/10.1037/cap0000216

Cambridge Dictionary. (n.d.). Expert. In *Cambridge University Press online dictionary*. Retrieved February 11, 2022, from https://dictionary.cambridge.org/dictionary/english/expert

Carter, J. A., Clark, A., Kallestrup, J., Palermos, S. O., & Pritchard, D. (Eds.). (2018a). *Extended epistemology*. Oxford University Press.

Carter, J. A., Clark, A., Kallestrup, J., Palermos, S. O., & Pritchard, D. (Eds.). (2018b). *Socially extended epistemology*. Oxford University Press.

Chambers, C. (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT Press.

PsychOpen GOLD

Clark, A. (2017). Embodied, situated, and distributed cognition. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 506–517). John Wiley & Sons. https://doi.org/10.1002/9781405164535.ch39

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Coady, D. (2012). *What to believe now: Applying epistemology to contemporary issues.* Wiley-Blackwell.

Code, L. (1984). Toward a 'responsibilist' epistemology. *Philosophy and Phenomenological Research, 45*(1), 29–50. https://doi.org/10.2307/2107325

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R. M. K., …Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour, 6*(12), 1731–1742. https://doi.org/10.1038/s41562-022-01458-9

Collins, H. (2018). Studies of expertise and experience. *Topoi, 37*(1), 67–77. https://doi.org/10.1007/s11245-016-9412-1

Collins, H., & Evans, R. (2008). *Rethinking expertise.* University of Chicago Press.

Crede, M., & Harms, P. (2019). Questionable research practices when using confirmatory factor analysis. *Journal of Managerial Psychology, 34*, 18–30. https://doi.org/10.1108/JMP-06-2018-0272

Dahdouh-Guebas, F., Ahimbisibwe, J., Van Moll, R., & Koedam, N. (2003). Neo-colonial science by the most industrialised upon the least developed countries in peer-reviewed publishing. *Scientometrics, 56*(3), 329–343. https://doi.org/10.1023/A:1022374703178

Derksen, M. (2017). A flair for science. *Mindwise.* https://mindwise-groningen.nl/a-flair-for-science/

de Ridder, J. (2014). Epistemic dependence and collective scientific knowledge. *Synthese, 191*(1), 37–53. https://doi.org/10.1007/s11229-013-0283-3

de Ridder, J. (2022). 'Three models for collective intellectual virtues.' In M. Alfano, C. Klein & J. de Ridder (Eds.), *Social virtue epistemology* (pp. 367–85). Routledge.

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., . . .Nosek, B. A. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology, 67*, 68–82. https://doi.org/10.1016/j.jesp.2015.10.012

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

Fontanarosa, P., Bauchner, H., & Flanagin A. (2017). Authorship and team science. *JAMA, 318*(24), 2433–2437. https://doi.org/10.1001/jama.2017.19341

Forscher, P. S., Basnight-Brown, D. M., Dutra, N., Adetula, A., Silan, M., & IJzerman, H. (2021). *Psychological science needs the entire globe, Part 3.* APS Observer, 35.

Forscher, P. S., Wagenmakers, E., Coles, N. A., Silan, M. A., Dutra, N. B., Basnight-Brown, D., & IJzerman, H. (2020). *The benefits, barriers, and risks of big team science.* PsyArXiv. https://doi.org/10.31234/osf.io/2mdxh

Frankenhuis, W., Panchanathan, K., & Smaldino, P. E. (2022, July 23). *Strategic ambiguity in the social sciences.* MetaArXiv. https://doi.org/10.31222/osf.io/kep5b

Fricker, E. (2006). Testimony and epistemic autonomy. In J. Lackey & E. Sosa (Eds.), *The epistemology of testimony* (pp. 225–250). Oxford University Press.

Fricker, M. (2012). Group testimony? The making of a collective good informant. *Philosophy and Phenomenological Research, 84*(2), 249–276. https://doi.org/10.1111/j.1933-1592.2011.00565.x

Fricker, M. (2020). Institutional epistemic vices: The case of inferential inertia. In I. J. Kidd, H. Battaly, & Q. Cassam (Eds.), *Vice epistemology* (pp. 89–107). Routledge.

Giere, R. N., & Moffatt, B. (2003). Distributed cognition: Where the cognitive and the social merge. *Social Studies of Science, 33*(2), 301–310. https://doi.org/10.1177/03063127030332017

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics, 33*(5), 587–606. https://doi.org/10.1016/j.socec.2004.09.033

Giner-Sorolla, R. (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science, 7*(6), 562–571. https://doi.org/10.1177/1745691612457576

Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research, 63*(1), 85–110. https://doi.org/10.1111/j.1933-1592.2001.tb00093.x

Goldman, A. (2014). Social process reliabilism: Solving justification problems in collective epistemology. In J. Lackey (Ed.), *Essays in collective epistemology* (online ed.). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199665792.003.0002

Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., & Ioannidis, J. P. (2022). Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). *Perspectives on Psychological Science, 17*(1), 239–251. https://doi.org/10.1177/1745691620979806

Hardwig, J. (1985). Epistemic dependence. *The Journal of Philosophy, 82*(7), 335–349. https://doi.org/10.2307/2026523

Hardwig, J. (1988). Evidence, testimony, and the problem of individualism—A response to Schmitt. *Social Epistemology, 2*(4), 309–321. https://doi.org/10.1080/02691728808578498

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29. https://doi.org/10.1038/466029a

Hurley, S. (1998). *Consciousness in action.* Harvard University Press.

Hutchins, E. (1995). *Cognition in the wild.* MIT Press.

IJzerman, H., Dutra, N., Silan, M., Adetula, A., Brown, D. M. B., & Forscher, P. (2021). Psychological science needs the entire globe, Part 1. *APS Observer,* 34–35.

Klein, R. A., Cook, C. L., Ebersole, C. R., Vitiello, C., Nosek, B. A., Chartier, C. R., Christopherson, C. D., Clay, S., Collisson, B., Crawford, J., Cromar, R., Vidamuerte, D., Gardiner, G., Gosnell, C., Grahe, J., Hall, C., Joy-Gaba, J., Legg, A. M., Levitan, C., …Ratliff, K. (2022). Many Labs 4: Failure

PsychOpen GOLD

to replicate mortality salience effect with and without original author involvement. *Collabra: Psychology, 8*(1), Article 35271. https://doi.org/10.1525/collabra.35271

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., ...Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443–490. https://doi.org/10.1177/2515245918810225

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science, 1*(3), 389–402. https://doi.org/10.1177/2515245918787489

Lehman, D. R., Chiu, C. Y., & Schaller, M. (2004). Psychology and culture. *Annual Review of Psychology, 55*, 689–714. https://doi.org/10.1146/annurev.psych.55.090902.141927

Levelt Committee. (2011, October 31). *Interim report regarding the breach of scientific integrity by Prof. D. A. Stapel.* Tilburg University, the Netherlands. Archived from the original on 27 June 2016. Retrieved on 15 February 2022 from https://web.archive.org/web/20160627142859/https:/www.tilburguniversity.edu/upload/547aa461-6cd1-48cd-801b-61c434a73f79_interim-report.pdf

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry.* Princeton University Press.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7*(6), 537–542. https://doi.org/10.1177/1745691612460688

Martin, B. L. (1973). Experts in policy processes: A contemporary perspective. *Polity 6*(2), 149–173. https://doi.org/10.2307/3234005

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

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science, 12*, 269–275. https://doi.org/10.1111/1467-9280.00350

Menary, R. (2006). Attacking the bounds of cognition. *Philosophical Psychology, 19*(3), 329–344. https://doi.org/10.1080/09515080600690557

Merriam-Webster Dictionary. (n.d.). Expert. In *Merriam-Webster.com dictionary.* Retrieved February 11, 2022, from https://www.merriam-webster.com/dictionary/expert

Mieg, H. A. (2006). Social and sociological factors in the development of expertise. In K. Anders Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 743–760). Cambridge University Press.

Montmarquet, J. A. (1987). Epistemic Virtue. *Mind, XCVI*(384), 482–497. https://doi.org/10.1093/mind/XCVI.384.482

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., …Chartier, C. R. (2018). The psychological science accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science, 1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*, Article 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73*, 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science, 7*(6), 615–631. https://doi.org/10.1177/1745691612459058

O'Donnell, M., Dev, A. S., Antonoplis, S., Baum, S. M., Benedetti, A. H., Brown, N. D., Carrillo, B., Choi, A. L., Connor, P., Donnelly, K., Ellwood-Lowe, M. E., Foushee, R., Jansen, R., Jarvis, S. N., Lundell-Creagh, R., Ocampo, J. M., Okafor, G. N., Azad, Z. R., Rosenblum, M., …Nelson, L. D. (2021). Empirical audit and review and an assessment of evidentiary value in research on the psychological consequences of scarcity. *Proceedings of the National Academy of Sciences of the United States of America, 118*(44), Article e2103313118. https://doi.org/10.1073/pnas.2103313118

Oude Maatman, F. (2021, July 12). *Psychology's theory crisis, and why formal modelling cannot solve it.* PsyArXiv. https://doi.org/10.31234/osf.io/puqvs

Oxford University Press. (n.d.). Expert. In *Oxford English Dictionary*. Retrieved February 11, 2022, from https://www.oed.com/view/Entry/66551?rskey=PeTuE5&result=1#eid

Palermos, S. O. (2022). Epistemic collaborations: Distributed cognition and virtue reliabilism. *Erkenntnis, 87*, 1481–1500. https://doi.org/10.1007/s10670-020-00258-9

Polanyi, M. (1966). *The tacit dimension*. Doubleday Anchor.

Popper, K. (2002a). *The logic of scientific discovery*. Routledge. (Original work published 1959)

Popper, K. (2002b). *Conjectures and refutations: The growth of scientific knowledge*. Routledge. (Original work published 1963)

Quast, C. (2018). Towards a balanced account of expertise. *Social Epistemology, 32*(6), 397–419. https://doi.org/10.1080/02691728.2018.1546349

Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. The University of Chicago Press.

PsychOpen GOLD

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2022). That's a lot to process! Pitfalls of popular path models. *Advances in Methods and Practices in Psychological Science, 5*(2). https://doi.org/10.1177/25152459221095827

Rosenstein, B. S., West, C. M., Bentzen, S. M., Alsner, J., Andreassen, C. N., Azria, D., Barnett, G. C., Baumann, M., Burnet, N., Chang-Claude, J., Chuang, E. Y., Coles, C. E., Dekker, A., De Ruyck, K., De Ruysscher, D., Drumea, K., Dunning, A. M., Easton, D., Eeles, R., …Radiogenomics Consortium. (2014). Radiogenomics: Radiobiology enters the era of big data and team science. *International Journal of Radiation Oncology, Biology, Physics, 89*(4), 709–713. https://doi.org/10.1016/j.ijrobp.2014.03.009

Rowlands, D. M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology, 22*(1), 1–19. https://doi.org/10.1080/09515080802703620

Rowlands, M. (1999). *The body in mind: Understanding cognitive processes*. Cambridge University Press.

Schimmack, U. (2022, January 3). Estimating the false discovery risk of psychology science. *Replicability Index*. https://replicationindex.com/2022/01/03/est-fdr-psy-sci/

Silan, M., Adetula, A., Basnight-Brown, D. M., Forscher, P. S., Dutra, N., & IJzerman, H. (2021, October 26). Psychological science needs the entire globe, Part 2. *APS Observer*.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366. https://doi.org/10.1177/0956797611417632

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science, 9*(1), 76–80. https://doi.org/10.1177/1745691613514755

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

Spellman, B. A. (2015). A short (personal) future history of Revolution 2.0. *Perspectives on Psychological Science, 10*, 886–899. https://doi.org/10.1177/1745691615609918

Tebes, J. K., Thai, N. D., & Matlin, S. L. (2014). Twenty-first century science as a relational process: From Eureka! to team science and a place for community psychology. *American Journal of Community Psychology, 53*, 475–490. https://doi.org/10.1007/s10464-014-9625-7

Tetlock, P. E. (2006). *Adversarial collaboration: Least feasible when most needed? Least needed when most feasible?* Presentation to board of directors of Russell Sage Foundation, New York City, NY, USA.

Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist, 76*(1), 116–129. https://doi.org/10.1037/amp0000622

Tomkins, A., Zhang, M., & Heavlin, W. D. (2017). Reviewer bias in single-versus double-blind peer review. *Proceedings of the National Academy of Sciences of the United States, 114*(48), 12708–12713. https://doi.org/10.1073/pnas.1707323114

Uhlmann, E. L., Ebersole, C. R., Chartier, C. R., Errington, T. M., Kidwell, M. C., Lai, C. K., McCarthy, R. J., Riegelman, A., Silberzahn, R., & Nosek, B. A. (2019). Scientific utopia III:

Crowdsourcing science. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 14*(5), 711–733. https://doi.org/10.1177/1745691619850561

Uygun Tunç, D. (2022). We should redefine scientific expertise: An extended virtue account. *European Journal for Philosophy of Science, 12*, Article 71. https://doi.org/10.1007/s13194-022-00498-2

Uygun Tunç, D. (2023). The subject of knowledge in collaborative science. *Synthese, 201*, Article 88. https://doi.org/10.1007/s11229-023-04080-y

Uygun Tunç, D., & Pritchard, D. (2022). *Collective epistemic vice in science: Lessons from the credibility crisis*. http://philsci-archive.pitt.edu/id/eprint/21120

Uygun Tunç, D., & Tunç, M. N. (2023). A falsificationist treatment of auxiliary hypotheses in social and behavioral sciences: Systematic replications framework. *PsyArXiV*. https://psyarxiv.com/pdm7y/

Vazire, S. (2017). Our obsession with eminence warps research. *Nature, 547*(7661), 7. https://doi.org/10.1038/547007a

Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science, 13*(4), 411–417. https://doi.org/10.1177/1745691617751884

Watson, J. C. (2022). *A history and philosophy of expertise: The nature and limits of authority*. Bloomsbury.

Weidman, A. C., Steckler, C. M., & Tracy, J. L. (2017). The jingle and jangle of emotion assessment: Imprecise measurement, casual scale usage, and conceptual fuzziness in emotion research. *Emotion, 17*(2), 267–295. https://doi.org/10.1037/emo0000226

Wilson, R. A. (2004). *Boundaries of the mind: The individual in the fragile sciences*. Cambridge University Press.

Winsberg, E., Huebner, B., & Kukla, R. (2014). Accountability and values in radically collaborative research. *Studies in History and Philosophy of Science Part A, 46*, 16–23. https://doi.org/10.1016/j.shpsa.2013.11.007

Zagzebski, L. T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.

Zwaan, R. A., Etz, A., Lucas, R., & Donnellan, M. (2018). Making replication mainstream. *Behavioral and Brain Sciences, 41*, Article e120. https://doi.org/10.1017/S0140525X17001972

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.