

FOCAL ARTICLE

Open science, closed doors: The perils and potential of open science for research in practice

Richard A. Guzzo^{1*}, Benjamin Schneider² , and Haig R. Nalbantian¹

¹Workforce Sciences Institute, Mercer and ²University of Maryland, Emeritus

*Corresponding author. Email: Rick.guzzo@mercer.com

(Received 19 March 2020; revised 10 March 2021; accepted 31 May 2021)

Abstract

This paper advocates for the value of open science in many areas of research. However, after briefly reviewing the fundamental principles underlying open science practices and their use and justification, the paper identifies four incompatibilities between those principles and scientific progress through applied research. The incompatibilities concern barriers to sharing and disclosure, limitations and deficiencies of overidentifying with hypothetico-deductive methods of inference, the paradox of replication efforts resulting in less robust findings, and changes to the professional research and publication culture such that it will narrow in favor of a specific style of research. Seven recommendations are presented to maximize the value of open science while minimizing its adverse effects on the advancement of science in practice.

There is a dramatic updraft of interest in the far-reaching concept of open science, and its pull is being felt throughout the organizational sciences. The movement is about making scientific methods, data, and findings more accessible and reaping the benefits that are presumed to follow.¹ It is also about broadening participation in the scientific enterprise such as through crowdsourced data collection (for example, iNaturalist.org is a curated site sponsored by the National Geographic Society and the California Academy of Sciences for compiling biodiversity observations) as well as encouraging research conducted by amateurs (e.g., “citizen science” such as described at The United Nations Educational, Scientific and Cultural Organization’s Global Open Access Portal, 2022; also see Göbel, 2019). Three overarching principles shape open science practices: transparency, sharing, and replication. These principles are germane to Grand, Rogelber, Banks et al.’s (2018) encompassing vision of “robust science” for all industrial-organizational (I-O) psychologists to rally around in the pursuit of increased scientific trustworthiness. That vision rests on other principles, too—rigor, relevance, theory, and cumulative knowledge—and it comes with prescriptions for actions that individuals and institutions can take to realize it. Our focus here is narrower, restricted to the three overarching open science principles. First, to provide the foundation for our major points, we briefly review open science principles and their implementation and supporting rationale (for more extended treatments see Banks et al., 2019; Grand, Rogelberg, Allen et al., 2018; Grand, Rogelberg, Banks et al., 2018; Toth et al., 2020). We then discuss in detail how acting on those principles may create negative consequences *for science* in disciplines that are founded on connectivity between science and practice. We close by offering a series of recommendations for ways of avoiding those undesirable consequences.

¹Open science is different from open access, though some open access journals feature themselves as open science (e.g., *The Journal of Personnel Assessment and Decisions* or *PAD*). *PAD* articles are available free to all for any legal use, but other principles of open science do not appear to be relevant (e.g., the requirement for transparency).

Open science principles in action

Let's use the structure of the traditional research journal article to illustrate how the principles of transparency, sharing, and replication move together and have their influence. After an introduction, the article's Methods section communicates the familiar essentials of who the research participants are, how research designs are executed, what stimulus materials are used, the nature of experimental manipulations, data collection processes, measures, and related matters. Available page space or journal conventions may limit the amount of detail that can be conveyed in the article itself, so the *transparency principle* calls for supplemental detail (of, for example, procedures including survey items) to be available via request and/or web access. The *sharing principle* calls for stimulus materials, instructions, measures, devices, and so on not just to be communicated but also to be made available so that others can use them in their own research, especially when attempting to reproduce results—the *replication principle*. Consistent with these principles, disclosing other relevant background resources such as computer code and statistical power calculations may also be encouraged or required as part of an article or in an available supplement.

The Results section of the traditional journal article then details the data analyses, including assumptions made, and reports the findings of those analyses. The open science model also calls for making available the original data for others to reanalyze.

In the spirit of sharing and transparency, open science calls for posting one's research process notes or one's "lab notes" that contain thoughts underlying research designs and hypotheses. A very specific open science tactic that is on the rise is preregistration. Preregistration involves, prior to conducting a research project, the public declaration of the hypotheses or questions to be addressed as well as descriptions of the measures, data collection methods, manipulations, sample size and characteristics, analytic techniques to be used, and other matters important to the planned research. These declarations/plans may become part of a journal's review process in which they are peer reviewed and accepted or rejected prior to data collection. Once accepted, a journal may commit to publishing a research report whatever results obtain or may impose a second round of manuscript review prior to committing to publication.

There is no shortage of resources to support acting on open science principles. Transparency checklists of conformation to open science principles are freely available; numerous sites exist to deposit and share data; and several preregistration sites exist, such as *aspredicted.org* hosted by the Wharton Credibility Lab, the American Economic Association's registry for randomized control trials at *socialscienceregistry.org* copyrighted to MIT, and *clinicaltrials.gov* for biomedical research. Table 1 illustrates a preregistration site (*aspredicted.org*). Noteworthy for what it implies about science–practice research in business organizations, this site does not accept preregistrations by researchers with email addresses ending in .com. Some preregistration sites are limited to recording experimental designs, and some of those also accommodate exploratory analyses as an adjunct to hypothesis-testing research. Visiting a few preregistration sites will quickly make apparent their emphasis on hypothetico-deductive approach to research.

In summary, guided by a few overarching principles and ample resources, open science offers an integrated set of practices that transform how the entire research process, from planning to methods to results, is conducted and made public through publication in refereed journal outlets.

Why open science?

"Efforts to promote open-science practices are, to a large extent, driven by a need to reduce questionable research practices" (Aguinis, Banks *et al.*, 2020, p. 27), and one of the major motivators for open science is fraud prevention. Fraudulent research not only hinders the advancement of knowledge but also tarnishes the reputation of science broadly, especially when news of fraud spreads through the general press. It is not difficult to find examples. *Science* retracted an article shortly after publication when it discovered that researchers had fabricated experimental results

Table 1 Example of Preregistration Questions

Author(s)
(name, email address, institution)
Submission
<ol style="list-style-type: none"> Data collection. Have any data been collected for this study already? Hypothesis. What's the main question being asked or hypothesis being tested in this study? Dependent variable. Describe the key dependent variable(s) specifying how they will be measured. Conditions. How many and which conditions will participants be assigned to? Analyses. Specify exactly which analyses you will conduct to examine the main question/hypothesis. Outliers and exclusions. Describe exactly how outliers will be defined and handled and your precise rule(s) for excluding observations. Sample size. How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about <i>exactly</i> how the number will be determined. Other. Anything else you would like to preregister? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?) Name. Give a title for this AsPredicted preregistration. Suggestion: use the name of the project, followed by study description.
<p>Finally, for record keeping purposes, please tell us the type of study you are preregistering.</p> <ul style="list-style-type: none"> • Class project or assignment • Experiment • Survey • Observational/archival study • Other _____

Notes: Preregistrations of research plans at this site are not permitted by (a) researchers whose email addresses end in.com or (b) answering Question 1 “yes.” Recommended responses for Questions 2–8 are “up to about 3,200 characters” each.

Source: <https://aspredicted.org/create.php> accessed 10 October 2020.

purporting to show that, in water, microplastics like those found in cosmetics make fish smaller, slower, and dumber (Berg, 2017). A university review board found 145 counts of fabrication or falsification of data in a body of research purporting to show that red wine consumption prolongs life (Wade, 2012). The American Medical Association retracted six papers in three journals coauthored by a social psychologist purporting to show that people eat more when food is served in bigger bowls and purchase higher calorie foods when shopping while hungry (Bauchner, 2018). Note that these examples touch on everyday behaviors of appearance and diet, suggesting that fraud at the intersection of science and practice may have significant reputational repercussions among broad audiences. Open science principles that help reduce fraud may thus be especially valuable to science–practice disciplines like I-O psychology where research fraud can have very important personal and organizational consequences.

The open science movement also seeks to overcome lesser vices, like *p*-hacking, which is selectively reporting results by analyzing and reanalyzing data in a search for a “*p*” value that is statistically significant. Several forces can motivate the exploration of data to settle on an analytic approach that yields desirably statistically significant findings. Whatever the motivations behind it, it is obvious how *p*-hacking is combatted through such tactics as publicly preregistering analysis plans and fully disclosing analytic details (e.g., how outliers were treated).

HARKing (hypothesizing after results are known) is another practice addressed by open science. Originally, HARKing was defined as presenting a post hoc hypothesis in a research report as though it had been formulated and tested prior to knowing the results (Kerr, 1998). This, of course, violates the tenets of the hypothetico-deductive approach to inference that underlies much—especially experimental—research, and in this original sense of the term HARKing would be considered a fraudulent misrepresentation. The concern with HARKing has evolved beyond the boundaries of fraud to include other after-the-data-are-in practices deemed to be questionable. Murphy and Aguinis (2019; Table 1) focus on two such practices. One is “cherry-picking,”

searching through measures or samples to find the results that support a particular hypothesis or research question; the other is “trolling,” searching through measures, interventions, or relationships to find something notable to write about. Through simulations involving changes in study parameters such as sample size and number of observed relationships, they find trolling to be the more pernicious.

There are serious discussions of HARKing’s benefits, such as by Kerr (1998), Hollenbeck and Wright (2017), and Vancouver (2018). In general, though, HARKing is widely viewed as an inferior way of learning from empirically verifiable observations because the researcher does not invoke a priori theories to predict a result. Indeed, those who discourage it seem committed to the idea that one conducts research to test a theory or set of theoretical principles and the project ends when the results are obtained to either falsify or validate the theory. As we will argue in a number of places later, this is a problematic underuse of talent, data, and thinking and is detrimental to scientific advancement in modern applied research.²

The principle of replication is so prominent in open science that the movement has become closely associated with efforts to resolve the “crisis of replication” in psychological research. Replicability has always been a concern for researchers, and the sources of variability in scientific findings are many, only a few of which relate to evils like fraud. Replication got the label “crisis” somewhat recently, emanating from reports of disappointing success rates of attempts to reproduce statistically significant findings in psychological research journals, reports that quickly reached audiences outside of scientific communities. For example, The Open Science Collaboration (2015) undertook an effort to reproduce the results of 100 published experimental and correlational studies in psychology and, judged against five criteria, declared about one third to one half of the replication efforts to be successful, results they found alarming. This promptly made its way to places like the front page of the *New York Times* (Carey, 2015). Not long thereafter, the journal *Nature* published the results of a survey of scientists on the “reproducibility crisis” (Baker & Penny, 2016). Respondents differed in their views of whether the crisis is “significant” or “slight” but, across multiple disciplines, scientists reported that the majority—often overwhelmingly so—of published work in their field is reproducible. A reanalysis of the Open Science Collaboration’s (2015) data was made by Gilbert *et al.* (2016), who found flaws in the replication efforts due to issues of sampling error, power, and infidelities between original studies and their replications. Gilbert *et al.*’s reanalysis estimated higher success rates of replications, about two thirds of the published studies, and demurred from declaring a crisis. Discussions that are specific to affairs in the organizational sciences mirror the broader discourse. For example, Cortina *et al.* (2017) endorse rigor through replication. Chen (2015, 2018) highlights the value of conceptual replications and seems less squarely in the “crisis” camp. Whether or not one sees a crisis, the open science movement is dedicated to increased replication.

Another attractive feature of open science is the enormous efficiency it creates. Sharing code and research materials makes the next researcher’s work that much easier; sharing data enables other researchers to do things that they might never have been able to do because of limited opportunity or means to gather original data. Time to publication shortens when journals commit to publishing results of studies with peer-reviewed research plans or preapprove the publication of replications, as has been suggested by some open science advocates. In addition, the discussion invited by the posting of lab notes, hypotheses, and designs might speed the process by which the originating researchers sharpen their thinking and elevate the quality of their research programs.

Open science has other ramifications as well, such as for training researchers and for interdisciplinary collaborations. It also is influencing journal publication practices. Some journals remain

²Labor/organizational economics, in contrast, emphasizes empirical tests of a priori mathematically expressed theoretical models, thus minimizing retroactive revisions of theory. However, such specificity comes with costs, including a narrowed focus of inquiry and the necessity of making numerous and sometime unrealistic assumptions to permit mathematical formulation and make models testable.

silent regarding open science practices, whereas others may encourage them, require them with exceptions, or require them without exception (e.g., data sharing). Open-access publication practices reflect the influence of open science. The Directory of Open Access Journals (2020) reports that 15,321 peer-reviewed, open-access journals exist. The journal *Judgment and Decision Making*, for example, will review the Introduction and Methods sections of preregistered studies and, if approved, will commit to publishing the finished studies whatever the results.³ Over 200 journals provisionally accept papers for publication based on peer review of such preregistrations (Center for Open Science, 2020a).

Open science incompatibilities with science–practice domains

We embrace the good that will come from the open science movement. However, open science is encumbered with incompatibilities between it and how much, if not most, of the research in science–practice disciplines is conducted. If not managed successfully, these incompatibilities will retard the advancement of knowledge and theory. One incompatibility is rather obvious: Reports of research conducted in practice domains often will be unable to meet open science requirements of disclosure and sharing for legitimate reasons related to risk, privacy, and intellectual property interests. The second incompatibility concerns theorizing. Open science practices clearly emphasize the hypothetico-deductive model of theory testing and creation and, consequently, tilt science–practice disciplines away from alternative ways of theorizing that are especially powerful in applied research. A third incompatibility is what we term the paradox of replication: The pursuit of replication in the open science model, we argue, will deliver less robust findings. Fourth, and most broadly, open science practices have the potential to create a narrowed research culture in which nonconforming studies are improperly devalued. Portents of some issues that we raise here can be found in papers such as those by Gabriel and Wessel (2013), Leavitt (2013), and Kepes and McDaniel (2013). We seek to provide an integrative account that is uniquely focused on a concern for scientific advancement through practice-oriented research. We elaborate each of these four points and then turn to potential solutions for the issues we raise.

Incompatibility #1: Disclosure and sharing

For research reports emanating from practice settings, the principles of transparency and sharing can become obstacles to their entry into the peer-reviewed literature. Data sharing is an especially problematic issue. Consider this example of research conducted in a fast-food business involving over 500 restaurants and over 20,000 employees. It used a prospective research design to investigate the effects of multiple current-period workforce and workplace attributes on three measures of next-period restaurant performance (profit, customer experience, speed of service) for each of 12 consecutive months (Erland et al., 2017). A large data set was amassed integrating observations from multiple sources. Some of the variables served as controls relevant to variance in performance measures such as the population density, median income levels, and unemployment rates of a restaurant's local community. These variables came from publicly available sources. However, most of the variables came from proprietary sources including the company's human resources information system (HRIS), financial, process control, and customer experience databases, much of which cannot be disclosed or shared. The company's competitive business advantages could be harmed by divulging fully transparent, quantitative details about the critical dependent variables, its restaurants' monthly performance metrics, or about certain performance-relevant covariates, such as the incidence of kitchen equipment repair. Employees could be exposed to risks of revealed identities if HRIS and workplace data are made available to others, and employers could be

³The commitment to publish results of preapproved research designs can be a way of remedying other faults such as publishing studies with low statistical power and a bias against publishing null results.

exposed to litigation risk to the extent that they are responsible for harm to employees through disclosure. Gabriel and Wessel (2013) discuss how data-sharing requirements can place vulnerable populations at risk and how such requirements can stop research from ever being initiated. Risks may be reduced by efforts to alter the data to make at least some of it shareable (e.g., tactics such as renaming, perturbing, and sharding). Banks *et al.* (2019) recognize the incompatibility between the principles of disclosure and sharing and research in organizations, and they suggest tactics to overcome the incompatibility such as delaying the release of original data and anonymization. However, none of the tactics mentioned here can guarantee that a clever data scientist could not identify at least some individuals through recombining identity-relevant variables in a data set (age, work location, job type, hours worked, hiring and termination information, etc.), especially if joined with local-area data from sources such as tweets, voter registries, and high school yearbooks. Adequate anonymization is an unresolved, multidisciplinary concern, and organizational researchers should be involved in the search for effective solutions.

Sometimes data from organizations will be fully shareable, and sometimes what is shareable is partial or is an abstraction of underlying data, such as a correlation matrix. Such a matrix may have variables deleted from it to avoid disclosures that are potentially harmful to an organization. Harm could come from reporting named variables that give strong identity clues or from reporting relationships, even among innocuously named variables, that are of value to competitors. Unfortunately, the greater the omission or obfuscation of variable names, the less valuable the shared matrix is for researchers' reanalysis or replication. Nonetheless, some data sharing is preferable to none. The *Journal of Business and Psychology*, for example, communicates a clear "research data policy" that encourages sharing of all materials and raw data described in a manuscript and makes compliance easy by directing authors to relevant data repositories and by providing scripted "data availability" statements for authors' to use. The journal's policy, although strongly pushing for sharing, acknowledges that not all data can be shared. An insistence on complete transparency and full data sharing is not an absolute requirement in widespread use today in order to publish. It is undeniably true, however, that there is movement in that direction.

Looking back, it is interesting to speculate how many landmark studies in the organizational sciences might have failed to "score high" if subjected to open science criteria. The influence of the Hawthorne Studies (Roethlisberger & Dickson, 1939), the AT&T management progress studies (Bray *et al.*, 1974), and the early performance appraisal studies at General Electric revealing how "split roles" for the appraiser can be useful (Meyer *et al.*, 1965) might never have registered.

The threat that we see from disclosure and sharing is to *hinder how frontline research informs science*. Banks *et al.* (2019), for example, observe that an organization may decline to participate in research that requires complete transparency and data sharing. We believe that imposing requirements of transparency and sharing such that they become obstacles to initiating research in organizations and barriers to the write-up of practice-based research reports for contribution to the scientific literature will have profoundly adverse consequences for the growth of realistic and useful knowledge (Rousseau, 2012).

It is not just the need to protect individuals' identity that is at issue. Additionally, companies cooperate in doing research such that organizations' identities are known to researchers so that relationships can be established between, for example, employee survey data and company-level performance data (e.g., Schneider *et al.*, 2003, on The Mayflower Group data). Companies will be appropriately loath to share such identifying information beyond the immediate team of researchers.

Furthermore, there is the important issue of intellectual property. Researchers in consulting firms and in-house research groups develop analytic tools and various selection, appraisal, and survey measures. Those organizations often may not wish to divulge the algorithms and/or measures to the public because of the investments they have made in their development. The research literature is replete with studies that have moved the field forward without such full disclosure and sharing being necessary. A recent paper by Pulakos *et al.* (2019) provides an example of how this

has typically been done but would no longer be possible if our journals required full disclosure and sharing. So, the Pulakos et al. article on agility and resilience in organizations provides a few sample items of the measure they developed and does not provide or make access available to the full measure or to the identity of the sample of organizations across which the new measure was validated successfully against organizational financial performance.

Lest we appear as mere rejectionists, each of this paper's authors has been involved in research efforts that could have complied with open science practices had those practices been in effect at the time. Examples include Guzzo and Waters (1982), Dieterly and Schneider (1974), and Nalbantian and Schotter (1995, 1997). We also have extensive records of research involving people at work in privately held companies, the locale of almost all of our data-based work. Under open science rules regarding disclosure and sharing, we would certainly have had far fewer opportunities to contribute to science through peer-reviewed publications than we have had to date. We estimate that the number of our data-based publications would actually be reduced by at least 75%. We would not be able to share data that came from HRIS systems in companies, where individuals were identified so that follow-up studies for panel analyses were possible (Schneider, White, et al., 1998), where proprietary data bases were used (the Center for Creative Leadership in Schneider, Smith, et al., 1998; A national insurance company's insurance agents and agencies in Schneider et al., 1996), or the study of the drivers of voluntary turnover at Fleet Bank based on the modeling of a panel data set constructed from the company's HRIS data and focused on actual stay/quit events (Nalbantian & Szostak, 2004), and so forth. The requirements for sharing and disclosure are but one issue that would inhibit the publication of the kind of research we do.

Incompatibility #2: Overidentification with the hypothetico-deductive model

Open science principles and practices are imbued with attributes of the classical hypothetico-deductive approach. Preregistration of hypotheses is a critical practice that promotes conformity with this way of building theory, as does the emphasis on replications of hypothesis tests. These types of open science practices explicitly rely on the hypothetico-deductive model of theory creation at a time when significant change is happening in the nature of science–practice research, especially research conducted in applied settings. That change involves “big data,” and it has big consequences for how best to advance theory.

Big-data organizational research is characterized both by lots of observations and by large numbers of variables. For example, Illingworth et al. (2016) describe selection research involving 125+ variables and a sample of 167,000 monthly assessments. They also describe research on coaching involving 80 to 120 variables. The previously mentioned Erland et al. (2017) case involved 78 variables in its analyses of the performance of 500+ restaurants and 135 variables in separate analyses of factors influencing voluntary turnover among 20,000+ employees. The latter is consistent with Hausknecht and Li's (2016) account of the multiplicity of variables and often-large samples in studies of turnover. The breadth of topics in I-O psychology addressable by large numbers of variables from multiple sources is described well by King et al. (2016).

The availability of big data in organizational settings enables “the accelerated development of existing theories to account for more relevant factors and to become more complex, specific, and nuanced—and ideally more accurate and useful as a result” (Guzzo, 2016, p. 348). A deficiency of the hypothetico-deductive approach is that it is, by definition, theory first. When 100 or more potentially meaningful variables are available for a research study in an organization of, say, 100,000 or more employees, there are many moderators, boundary conditions, levels of analysis, and causal pathways that are open to empirical tests. However, existing theories have not anticipated—could not be expected to have anticipated—such complexity and nuance. Consequently, existing theories offer too few hypotheses relative to the number of potentially meaningful relationships that can be rigorously tested in big-data research.

The emphasis on theory often leads to characterizing research as being of either two types, confirmatory or exploratory (Nosek *et al.*, 2018, speak of prediction or postdiction), the former being aligned with theory testing.⁴ Big-data research exposes the limitations of this distinction. The absence of sufficiently complex theories in a big-data world makes thoroughly hypothesis-driven prediction impossible, thus “confirmatory” is impaired as a referent category. “Exploratory” is an old term loaded with connotations of being the poor relative of confirmatory research. Furthermore, the label obscures the value of alternative, orderly approaches to theory generation and testing such as induction and abduction, discussed below. The presence of big data also breaks the confirmatory–exploratory mold by creating within-study opportunities to advance and continuously test and retest explanations, say, with randomly selected samples of 10,000 or more from a larger data set—essentially enabling multiple replications and/or meta-analysis processes within a single study’s large data set. McShane and Böckenholt (2017) present examples of single-paper meta-analysis in consumer behavior research. Within-study multiple replications and meta-analyses that are made possible with big data are at this time admittedly more aspirational than realized in publications. Indicative, though, is the work of Guzzo *et al.* (2014), which reports results from 34 distinct replications, involving nearly a million employees, of investigations of the effects of various forms of compensation on voluntary turnover. Also indicative is Guzzo *et al.*’s (2022) single-paper meta-analysis of 23 studies, each averaging over 40 measured variables, investigating the influence of employee age and tenure on work unit performance. Existing theories and processes of discovery are concurrently embedded in those works. Indeed, with the proliferation of continuously refreshed and expanded data in organizations, the empirical testing that becomes possible constitutes an approach to theory building beyond anything achievable with traditional methods tied to the hypothetico-deductive approach. Practices that elevate traditional hypothetico-deductive approaches over alternatives that are better suited to big-data realities will hold back the advancement of knowledge. What are such alternatives?

Some research scholars propose that an inductive model where one explores interesting questions and obtains significant unexpected findings for later pursuit is a superior process for theorizing (Locke, 2007). Locke illustrates how inductive processes based on research in both laboratory and field settings fueled the development of goal-setting theory, including the identification of moderating conditions that influence goals’ effects. The inductive model, indeed, may be quite useful for research in applied settings where questions emerge and data are available to explore possible answers to those questions, especially in our contemporary world that permits repeated testing of possible answers (cross-validation in I-O terms) on large data sets. McAbee *et al.* (2016) also assert the value of inductive research, arguing that the organizational sciences would benefit from a “shock” to rejuvenate empirical advances and that big-data analytics in organizations joined with inductive inference deliver that needed shock. They also make it clear that an embrace of big-data analytics does not imply abandonment of foundational principles of measurement and causal inference.

It is safe to say that much of the work done by practitioner-applied psychologists has not been of the hypothetico-deductive model but of the inductive model, which, in our experience, proceeds from a question not a theory. For example, one of our streams of research on service climate (Schneider, 1973; Schneider *et al.*, 1980) began with this question raised based on the 1973 paper: if customers in a bank’s branches say that they get lousy service because the employees are not service oriented, I wonder what kinds of conditions could be created that would encourage them to deliver superior service? This is how the stream of research on service climate began (see Hong *et al.*, 2013, for a meta-analysis of this line of research).

⁴Indeed, open science privileges confirmatory research, as evidenced by preregistration sites’ emphasis on hypothesis testing, efforts to discourage “questionable practices” that involve creating new explanations after data are collected, and the value placed on replication studies.

Another model for theorizing is abduction, sometimes referred to as “inference to the best explanation” (Douven, 2017). Unlike deductive or inductive inference making, abduction is about reaching credible explanations based on ensembles of facts, often involving multistage data analysis to arrive at the best of competing explanations (Haig, 2005). It is especially well suited to big-data research. The authors’ research-based consulting often embodies the spirit of abduction. For example, when studying internal labor market dynamics (Nalbantian et al., 2004), investigation is made by statistically modeling drivers of a set of outcomes: turnover, promotion, internal mobility, performance, and pay. Then the constellation of results is integrated to determine which of a set of preexisting theories relevant to these outcomes, in combination with empirical uniquenesses of the studied organization, best explains the patterns observed. Deploying theories this way helps uncover the story in the data in a compelling way that is easy to communicate to decision makers and easy for them to act on. It is also a more practical way of connecting theories to practice than starting with the presumption of one particular theory or model and then just testing it alone.

Prosperi et al. (2019) discuss the value of the joint presence of abductive and deductive processes in such research. They also directly challenge the wisdom of prevailing precautions against HARKing that are part of the open science movement, asserting that they conflict with the “modern reality” of researchers’ “obligations” to use big data (Prosperi et al., 2019, p. 1). Behfar and Okhuysen (2018) also assert that abduction is a needed complement to hypothetico-deductive reasoning in the organizational sciences. Heckman and Singer (2017) make a case for abduction in economics, saying that it helps overcome shortfalls in doctrinal theories, and it challenges prevailing schools of thought that are unable to account for new and surprising facts. The abduction approach also is naturally more aligned with the systems view of organizations, a point we elaborate below, and opens the door to new theorizing in ways that traditional theory-driven testing would not encourage.

In the next section we will discuss the paradox of replication where we will note that qualitative case methods represent the antithesis of the requirement for replication (Pratt et al., 2019). Here we note that such methods inherently begin with the goal of discovery not validation, the opposite of the hypothetico-deductive model. Indeed, projects in single organizations, by either qualitative methods or the use of a variety of sources for quantitative data, can both be thought of as case studies: Explorations of data from a single organization with no or few prior hypotheses are as much a case study as though it were done qualitatively. In other words, a primary open science principle devalues two essential streams of research in organizational psychology, potentially rendering them inappropriate for subsequent publication: qualitative case methods and quantitative case methods.

A bottom line to our thinking on the issue of theory and theorizing is that (a) our field has simply overemphasized the necessity for having a theory to test before one begins research and that (b) the open science movement concretizes this overemphasis. In the early 2000s the *Journal of Applied Psychology* (JAP), felt compelled to issue a call for theory-driven research. The editors of the JAP special issue on theory (Klein & Zedeck, 2004) put the need for theory this way: “Research in the absence of theory is often trivial—a technical feat more likely to yield confusion and boredom than insight. In contrast, research that is guided by theory, or that develops theory, generates understanding and excitement” (p. 931). It follows that problem-focused, inductively oriented work is trivial and of course uninteresting to JAP—at least in 2004.

Economics research is theory driven far more than problem driven. During the mid- and late-20th century, important theoretical developments appeared such as human capital theory, the operation of internal labor markets, and personnel economics. Unfortunately, the practical application and demonstrated value of these developments have fallen short of their theoretical promise. A major reason for this is inadequate access to organizational data and the resulting inability to inform theory by research that deals with practical issues. A recourse to laboratory experimentation ensued to test theories of incentives, decision making, and pricing mechanisms in organizations. As one of the authors (Nalbantian) can attest from personal research experience,

experimental economics offers powerful insights on narrowly defined, technical questions. However, it does not come close to generating the learning that arises from rigorous examination of real-world data, which reflect actual decisions and outcomes in organizations and which create opportunities to examine whether theoretical predictions hold up across a multitude of contexts and circumstances.

In 2008, a wonderful review of the research published in *JAP* and *Personnel Psychology* by Cascio and Aguinis (2008) revealed that (a) practitioners were not publishing in these journals (*JAP* had 4.67% of its articles by practitioners and (b) as would be expected from this figure, the articles were of not much relevance to the world of practice. Cascio and Aguinis concluded their review this way:

On the basis of our review, if we extrapolate past emphases in published research to the next 10 years, we are confronted with one compelling conclusion, namely, that I–O psychology will not be out front in influencing the debate on issues that are (or will be) of broad organizational and societal appeal. It will not produce a substantial body of research that will inform HR practitioners, senior managers, or outside stakeholders, such as funding agencies, public policymakers (including elected officials), or university administrators who control budgets. (p. 1074)

Early in our field, questions of import to individual and organizational performance drove applied research, but since then such important practice issues have given way to academic research driven by theory testing. The latter fits well with the open science model but the former fits better with the way research is done in practice settings. We are not alone in our views that theory is not the be all and end all for guiding research. Campbell and Wilmont (2018) argue that there are fundamental flaws in the theory-building process in the organizational sciences, including an overemphasis on theory for theory's sake. Critiques of obsession with theory predate the rise of open science, such as that offered by Hambrick (2007) on the role of theory in management research:

I suspect that many members of our field, including those in leadership positions, believe that our hyper commitment to theory—and particularly the requirement that every article must contribute to theory—is somehow on the side of the angels. They may believe that this is a hallmark of a serious field. They may believe that theory is good and that the “mere” description of phenomena and generation of facts are bad. Worse yet, they may have given no thought to these matters, accepting our field's zeal about theory as simply part of the cosmos. (p. 1351)

Of course, we must make it clear that we are definitely in favor of deductive theory writing and research and have engaged in such efforts ourselves (Heilman & Guzzo, 1978; Nalbantian & Schotter, 1997; Schneider, 1975; 1983). What we are not in favor of is making hypothetico-deductive theory writing and testing the predominant way to get articles published in the refereed journal space. The open science model, with its emphasis on hypothesis testing based on pre-existing theories, can very easily become an impediment to knowledge advancement through problem-focused inductive and abductive work.

Incompatibility #3: The paradox of replication

The emphasis on theory-driven work along with disclosure and sharing create a serious paradox with respect to establishing robust relationships. Under pressures to replicate, what type of replication studies do we think researchers will attempt most frequently? It will be studies with simple hypotheses tested on small samples with few variables and not situated in applied settings. This

type of study is most able to comply fully with open science principles designed to foster replication, and publication practices based on those principles will encourage more of the same. This is not the best way to improve the robustness of a field's research findings. By comparison, complex questions studied on very large data sets can markedly improve the reliability of findings because such research can test relationships under a variety of circumstances and explicitly control for relevant variables (boundary conditions) that offer potentially competing hypotheses. In addition of course and as noted earlier, repeated randomly chosen data sets from very large databases can be used to do replications immediately, providing distributions of findings and relationships then testable via meta-analytic type models.

But small sample research prevails. Hernandez et al. (2016) reported a median sample size of 73 for articles in *Psychological Science*, 90 for the *Journal of Personality and Social Psychology*, and 80 for regular articles (52 for brief articles) in *Cognition*. Marszalek et al. (2011) tracked four journals over time (*Journal of Applied Psychology* among them) and found median sample sizes of 48, 32, and 40, respectively, in 1977, 1995, and 2006. Shen et al. (2011) tracked *JAP* from 1995 to 2008 and reported median sample sizes of 173 for studies of individuals and 65 for studies of groups, and K uhberger et al. (2014) randomly sampled 447 peer-reviewed journal articles in psychology and found that 85% had sample sizes less than 200. Small samples prevail in economics research as well. Ioannidis et al. (2017) quantitatively examined 159 meta-analyses covering 6,700 distinct empirical studies across a range of topic areas and found that in 50% of the areas studied, about 90% of the studies were underpowered. Further, more than 20% of the areas lacked even a single study with adequate statistical power. These results show that economic research, perhaps like organizational science research broadly, is extensively afflicted by residual bias that undermines the credibility—and likely replicability—of the statistically significant results reported.

Open science's emphasis on replication to attempt to assure the certainty of published findings will encourage projects with small sample sizes that should necessarily raise concerns of adequate statistical power. But, when it comes to theory development, small-sample research has major implications beyond statistical power. It follows from small samples that (a) comparatively few variables can be analyzed in any one study, (b) few tests are able to be made of theoretical propositions of moderators or boundary conditions, and (c) there is limited opportunity to create new theoretical insights based on multivariate fact patterns. That is, the typical small to medium sample sizes in theory-testing research are actually incompatible with the need for strong tests of theory to achieve robust findings. Big-data efforts that are possible in practice permit, through induction and abduction, the testing and elaboration of complex theories based on robust findings. Proserpi et al. (2019) offer a new way of thinking about replication with big data. They suggest that it is usefully thought of as providing opportunities to reproduce (replicate) multivariate models of fact patterns rather than a very specific and narrow observation or inference.

Furthermore, there are blind spots in open-science-driven advocacy of replication. One is an overvaluing of exact replication of specific findings and an undervaluing of the role of meta-analysis of diverse studies as a means of establishing the robustness of relationships. Exact replication primarily is based on the degree to which a study finds or fails to find statistical significance. The problem is that we know, or should know, that statistical significance is highly dependent on sample size. It follows that with small sample sizes it is *expected* that there will be variability in the degree to which a given study will have its statistical significance replicated. It has been almost 45 years since Schmidt and Hunter (1977) began showing that small sample sizes for studies masked the likely true relationships that were observed. In a summary of this work, Schmidt (2010) makes the following appropriate observation: "researchers appear to be virtually addicted to significance testing, and they hold many false beliefs about significance tests (p. 237)." The major false belief is that statistical significance indicates the probability of a study being replicated! Another blind spot concerns the association between sample size and effect size. K uhberger et al. (2014) report a $-.54$ correlation between sample size and effect size. A literature of replicated small-sample studies thus runs the risk of overestimating the magnitude of "true" relationships. It is our personal experience

that big-data-style research indeed yields smaller effect size estimates of X–Y relationships when compared with reports of the same X–Y relationship in published literature precisely because the big-data-style research controls for many other reasonable sources of variance in X and Y that the published research did not.

The paradox, then, is that attempts at replicating small-sample research will lead us away from robust theoretical insights. Although research conducted in practice settings may not be able to comply with the conditions that facilitate exact replication—full disclosure, sharing—and thus may be less likely to be replicated exactly, compared with its small-sample counterparts big-data-style research can yield more robust findings along with less risk of effect size overestimation.

To close the loop on the paradox of replication, we noted earlier that there is no logic that would require replication of qualitative case studies. As Pratt *et al.* (2019) noted in their *Administrative Science Quarterly* editorial essay, “replication misses the point of what the [case study] work is meant to accomplish” (p. 1). They note that the emphasis on replication in open science is an attempt to achieve trustworthiness but that “management journals need to tackle the core issues raised by the tumult over transparency by identifying solutions for enhanced trustworthiness that recognize the unique strengths and considerations of different methodological approaches in our field” (p. 1). Later we present some proposals for how to help resolve some of these replication issues with respect to publication.

Incompatibility #4: Evolving cultural and professional norms

We, like many readers, grew up in a professional milieu characterized by specific norms about how to conduct and publish research. Many of the old challenges, tactics, and rhythms of research that were part of that milieu have disappeared, thankfully. Mailing paper surveys, carrying boxes of IBM cards containing punched-in data to the computer center and waiting for printouts from behind the wall (and despairing at their thinness due to an error message), the snail’s pace of manuscript preparation, submission, and dissemination are now absent. Web-based surveys, ample computing power, and electronic management of manuscripts greatly improve the processes of conducting and publishing research. Important positive norms have persisted through these changes, such as the value of peer review in the publication process. However, we ask the following question: What happens when other norms that define “good” research change? Our fear is that practices aligned with open science principles will result in dramatic alterations in the norms that define “good” research in ways that may be bad for science–practice disciplines and what actually gets published in our journals and disseminated to the broadest audiences. We see several worrisome signs—preregistration sites that preclude use by business-based researchers, the privileging of confirmatory research, the presumption that applying open science principles in journals’ peer review processes is of interest only to academics and not to practitioners (Society for Industrial and Organizational Psychology, 2020)—that a cultural shift may be underway that devalues research that cannot fully conform to open-science-driven ideals. That cultural shift will unnecessarily (a) deter scholars from accessing the research opportunities presented by the explosion of useful data in organizations and (b) reduce the opportunities of practice-based researchers to contribute to the body of scientific literature. Another aspect of concern is the risk that open science practices will result in methodological sameness.

Consider the incompatibilities we have discussed and their potential effects on the kind of research likely to emerge. We have made our prediction about the type of replication research that will emerge. But what about other types of research? If disclosure and sharing become requirements for research to be deemed good and appropriate, where does that leave researchers who use proprietary measures, data, and samples that cannot be shared and/or disclosed (King & Persily, 2019)? Will we then be a science of simple questions studied on small samples yielding unreliable and findings that are likely irrelevant to what happens in real organizations? If research is only deemed to be good when the hypotheses and the methods are stated in advance, what does

that do to researchers who ask interesting and important practical questions and find new and interesting ways of developing measures (like via text analytics, for example) as the research process unfolds? What will happen to big, bold research endeavors that probe complex interactions of factors in and across organizational systems? Will our research become characterized by the laboratory studies now becoming so prominent in business schools that they hire more social psychologists so they can be “really” scientific? As Barry Staw (2016) has recently noted in his essay reflecting on trends he has observed in the field, “Although an increasing number of social psychologists are now housed in business schools, they are still doing social psychological rather than organizational research” (pp. 16–17). Staw’s Table 1 (p. 14) is very instructive wherein he notes that the social psychologist in business schools pursues lab and field experimentation to search for underlying processes rather than using multiple methods (including qualitative methods, observational and historical data) to search for upward, outward, and downward influences. Is small-sample, theory-first research the be all and end all of a science that has practical relevance? We see a real risk that professional norms overly aligned with open science principles will evolve the culture too far in this direction.

There is increasing decrying of the failure to use the newer sources of data available to both academics and practitioners. The paradox is that there has been a reversion to simpler studies with small samples, as Staw (2016) describes, limiting drastically (a) what appears in the peer-reviewed literature and (b) the relevance to organizational life. At least two factors have yielded this outcome, an outcome being pushed further by the open science movement. First, as King and Persily (2019) note in their prescient article on industry–academic partnerships in research, the existing culture for academic researchers has been characterized by the independent scholar who had “unfettered control over their research agenda, methodological choices, and publication options” (p. 1). They note that in the era of big data (they write as political scientists) that this traditional model no longer works if research is to be relevant to the real world. They suggest, as do we, that unless researchers have access to the vast big-data sources that exist, they will increasingly focus on smaller, societally irrelevant issues. The second factor paradoxically pushing the hypothetico-deductive model on smaller samples with fewer variables is the reward system that exists in the academic publishing world where publications in what have come to be called “A” outlets are the only ones that count for promotion. The perfect article is the one that generates the fewest methodological and conceptual issues that can be attacked by reviewers, so the surest way to get a publication is to conduct a study with clear hypotheses using standard, favored methodologies and having as much control as possible over the research process and outcomes. Aguinis, Cummings, et al. (2020) thoroughly document the A-journal fetish, worldwide, noting that “there are mounting concerns about unintended negative effects of using A-journal lists to assess research value. Among these deleterious outcomes are questionable research practices, narrowing of research topics, theories, and methods, and lessening research care” (pp. 136–137). Sound familiar? There appears to be an implicit conspiracy fighting against research that is relevant to organizational life. Aguinis, Cummings, et al. reach the strange conclusion that this is not as bad as one might think with respect to the information business leaders have as a basis for making decisions. That is, they note that textbooks and articles in the press use the materials published by consulting firms and companies as sources for insights, not research published in academic journals. In other words, to have information that is relevant to the world of practice, textbook authors and journalists *no longer rely on the research published in academic journals*. Is this the world of evidence that open science initiatives wish to produce, one in which students and the public are being taught about important issues based on reports and results that have not met usual and typical peer-reviewed research standards?

Moreover, we believe that enthusiasm for open science runs the risk of creating overbearing methodological fashion. Such is evident in economics, where methodological conformity is fully on display in empirical research and that has led to greater concern with the wizardry of technique than with the value and practical relevance of insights generated from data. Labor/organizational

economics has long been concerned with understanding phenomena like labor market participation, wage determination, the employment effects of minimum wage laws, human capital development, and the value of education and training, among other topics. Empirical analyses tend to draw on Census data and a relatively few other large national data bases such as the Panel Study of Income Dynamics and National Longitudinal Survey of Youth, surveys that track large cohorts of individuals over time. Scholars repeatedly going to the same well often find themselves relying on econometric dexterity to uncover something new. Another aspect of methodological uniformity in labor and organizational economics is an overreliance on a single type of model specification when making empirical tests, fixed effects models. The culture of conformity to this single “standard” foregoes the benefits of alternative model specifications that take better advantage of the value of big data for both theoretical and practical insights.

Recommendations

As is often true in life, good comes with bad. This paper’s assessment of the open science movement emphasizes its pitfalls and shortcomings that we believe are especially relevant to science–practice disciplines and the research that gets done and published. As we stated previously, good will result from open science principles but not without some modifications of current and developing practices tied to those principles. Here we offer seven specific recommendations designed to preserve benefits of open science practices while minimizing their unwanted negative consequences.

Encourage more collaboration between academic- and practice-based researchers

Grand, Rogelberg, Allen *et al.* (2018) offer this recommendation, and we repeat it here for its importance. In addition to the value inherent in combining the two perspectives when designing, analyzing, and interpreting studies, academic-based researchers are likely to see to it that practice-based researchers adopt open science practices wherever possible and are more likely to connect existing theories to research done with “real” databases. More and better publications of practice-based research will result. Such collaborations will foster a constructive culture of science–practice research. In the continued absence of such collaborations, the fault lines between “academic” and “applied” could easily widen because of one side’s advocacy of open science practices that the other side cannot satisfy.

This call for more collaboration is not just the echo of past exhortations. It is a renewal, a response to the center of gravity of data now found in organizations (Guzzo, 2011). Organizations and their vendors have large volumes of relevant data within easy reach. Sources include HRIS systems (variables related to individual employees’ performance, jobs, experiences, leaders, coworkers, and workplace characteristics), learning management systems (providing records of training and development opportunities taken), communication (e.g., email, SMS records) databases documenting workplace interactions, applicant databases with such data as employment screening and work history variables, employee engagement databases, benefits elections and utilization databases that provide insights into such things as employees’ preferences and their well-being, and the records of routinely collected business outcomes (e.g., sales, customer retention). A great many of these variables take the very same form as they would had they been created specifically for research purposes (e.g., performance ratings in organizational databases often look just like performance ratings created by researchers). Such data, along with the presence of researchers in organizations, are good reasons for collaborating for the sake of better science. Organizational sciences risk being bypassed in relevance if they miss the opportunities being created by the proliferation of digital data on so many aspects of behavior, processes, and contexts.

Adapt the peer-review process to meet objectives of the open science transparency principle but without placing individuals or organizations at risk

Our focus here is on research reports from naturally occurring organizational settings where legal or other prohibitions prevent full disclosure to protect privacy or proprietary interests. The objectives of an adapted review process are to (a) pull back the veil so that a trusted outside party or parties can verify the authenticity of the details of the reported research effort and (b) remove one of the barriers (full disclosure) to publishing. One approach might be to provide a special reviewer (chosen from among acknowledged experts who would be uninvolved in decisions about the specific manuscript), after they sign a binding nondisclosure agreement (NDA), the right to see the original data and to discuss with the researchers details of methods of analysis and other relevant considerations. Checklists and journal policy standards could aid the process. This style of review would protect research integrity and provide a valuable imprimatur for the consumers of published research. Other approaches may exist. The process suggested here is not a review of the proposed article but a review of the methods/procedures/sample, and it does not preclude other forms of blind reviews of the same research report with respect to the possible actual publication value of the research effort itself.

As we prepared this article, and this recommendation, we became aware of a similar proposal for research in political science. King and Persily (2019) demonstrate how big data can be so facilitative of the research done in their field but lament the limitations to accessing such data, limitations like those in our field. As they note, “Today, big data collected by firms about individuals and human societies is more informative than ever, which means it has increasing scientific value but also more potential to violate individual privacy or help a firm’s competitors” (p. 2). Their proposed solution to the dilemma of access and privacy is to create two interacting but independent groups, one of academics who wish to do the research and a second of trusted third parties who sign NDAs regarding the details about a firm and its database. The third party “thus provides a public service by certifying to the academic community the legitimacy of any data provided.”

In short, there are tactics available to ensure that the research process and the research data meet the highest possible standards without giving rise to the risks that compliance with open science practices would create. We are also hopeful that journal reviewers are open to the alternatives to hypothetico-deductive models with big-data research reports. Needless to say, processes ensuring the privacy of data have the potential to make the most important data available to academic researchers, thus encouraging the best research on the most relevant possible databases for publication to other academics, to executives who require validated evidence, and to textbook authors and news reporters as well.

Revise how we teach theorizing

Training in the hypothetico-deductive approach is well rooted and closely integrated with training in inferential statistical methods. The advent of big data requires that we add mastery of inductive and abductive methods of theory development to that foundation. As a discipline, we have caught on to the data analysis innovations that come with big data. For example, see the “scenic tour” of statistical methods for big data offered by Oswald and Putka (2016) and the more recent review by Oswald et al. (2020). Although we may be quick at taking up these new methods, we see a need to increase our collective ability to advance theory from their application to our big-data research. This is where induction and abduction fill the void left by hypothetico-deductive inference. Both of these alternative approaches concern the process of testing and building theory from the integration of multiple findings and fact patterns illuminated by “big models” that facilitate replication and validation, making data analysis an opportunity to learn the unanticipated in addition to testing preexisting theory and, not so incidentally, allow for replication and meta-analysis of findings at the same time.

Do not specifically incentivize the publication of replications, especially exact (“direct”) replications

Journals can incentivize replications by inviting them, by creating special issues devoted to them, and by guaranteeing the publication of results of preregistered, peer-reviewed replication research plans. But if a replication crisis exists, such tactics are likely to magnify it. As we have discussed, attempts to replicate only a subset of extant research reports will predominate: the easy-to-do studies with smaller samples and fewer variables that are free of the constraints encountered in applied settings.⁵ As the Center for Open Science recognizes, “journals and [professional] societies shape the incentives and reward structures driving researchers’ behavior” (Center for Open Science, 2020b). Incentivizing replications with the rewards of publication will shape behavior and the research literature in undesirable ways for science–practice disciplines, limit what we learn, limit the variety of methods we will use, and narrow the scope of what defines “good” research (Aguinis, Cummings *et al.*, 2020).

Embrace conceptual replications and encourage them when big-data sets are available for exploring contingencies

Conceptual replications address previously tested theories and propositions but do so by being different from prior research “in the operationalizations of the phenomenon, the independent and dependent variables, the type and design of the study, and the participant populations” (Crandall & Sherman, 2015, p. 93). Crandall and Sherman make the case that conceptual replications deliver far more scientific value than do direct replications. A primary reason for their superiority concerns what is learned when replication studies fail to recreate prior findings. When the usual exact replication fails, it is impossible to discern why: Was it because the theoretical concept is poorly specified or the operationalizations are faulty or because, as in the original study, the sample was too small? Exact replications—no matter how many times they are attempted—do not resolve this ambiguity. Conceptual replications can productively disperse sources of failure to replicate across operationalizations, settings, samples, and so on. Consequently, conceptual replications are far superior for establishing generalizability and identifying why and when theorized relationships hold. In other words, conceptual replications are the preferred way of creating robust research results and theory.

Conceptual replication opportunities abound in big-data research precisely because it enables testing of previously reported findings in settings different from prior studies, with different operationalizations of variables, alternative methods of data collection and analysis, with different participants, with more covariates and boundary conditions examined, and so on. The key is to get reports of such big-data research published. This way a body of robust findings—and replications—will grow quickly, without special inducements, in science–practice domains. Big-data research need not originate only in applied settings; it can occur in simulations and controlled experiments, too. However, we estimate that applied settings are the more likely source of higher volumes of such research.

Affirm the value of preregistration, where appropriate

Preregistration of research plans—including designs, measures, and analyses—is a valuable tool for combatting research ills such as *p*-hacking and for promoting open discourse early in a research project. However, it is unworkable for many research efforts, especially for research in applied settings and not only for inductive or abductive work. That is, although still rare, there

⁵Not all replication efforts are small scale. The Center for Open Science in partnership with the Defense Advanced Research Projects Agency is sponsoring a multicollaborator investigation involving replications and reanalyses of numerous published studies. See <https://www.cos.io/score>.

is a growing literature of field experiments (Eden, 2017) contributing to our research. On the surface, such field experiments would seem to fulfill open science principles given the focus of them on specific issues. However, companies may be averse to revealing who they are, who participated in the research, and which proprietary measures and procedures are used. Also, it is truly unrealistic to require researchers such as Illingworth et al. (2016) to fully preregister the over 125 variables used in their selection research or the 135 variables of Erland et al.'s (2017) analysis of turnover. Aguinis, Banks et al. (2020), publishing in a journal with a primarily academic rather than practitioner audience, state that it requires "30–60 min of authors' time to preregister" (Table 1; p. 29). This is out of touch with the realities of multivariate, big-data research in applied settings. Beyond the mere burden of describing what could be well over a hundred variables, a task that constitutes only one small portion of what is called for in preregistration, research in applied settings such as business organizations only sometimes can be fully preplanned. By necessity, such research efforts, especially when oriented to finding solutions to pressing problems, often must respond to changing circumstances, take advantage of data or samples that become available mid-stream, or change tactics to respond to restatements of the problem driven by the results of initial analyses. These points are fairly straightforward and evident, but they can result in finished research projects that look quite different from initial plans. The concern we have is that in the advocacy of preregistration as a tactic for all there is too little detailed explication of when the tactic is ill-suited and when its enforcement will discourage very good practice-focused research from entering the research literature. Give credit for preregistering a study when it is useful and supportive of scientific advancement, but do not discredit research when it is appropriately not preregistered. These latter cases, of course, would benefit from our previous argument about having NDAs for reviewers and/or a panel of senior academics to review methods and analyses without the necessity of researchers and organizations bearing the risks of full disclosure.

Preregistration as a tactic impinges on the true nature of scientific advancement and the role of unexpected discoveries, especially when studying social systems such as organizations. This is nicely illustrated by the distinction in economics research between partial and general equilibrium contexts. Theory-driven research works best in what economists call a partial equilibrium context. In that context, narrowly defined relationships that can be easily expressed mathematically and isolated in the data are tested. It is decidedly not good for addressing general equilibrium outcomes, those that reflect the simultaneous operation of complex interactions among system components. It is very hard to prespecify theoretical propositions about those interactions where often there are inflection points in specific system components and nonlinearities that can produce dramatic changes in the characteristics of the system as a whole and of the behavior of individuals and groups in those systems. Where those inflection points and nonlinearities come into play is an empirical question neither easily predicted from theory nor easily described in advance. From this perspective, then, current preregistration practices appear more to serve the advancement of marginal findings than of transformational findings.

Affirm the value of sharing, where appropriate

Sharing of data, code, methods of analysis, materials, and so forth, like preregistration, is a valuable tactic but as noted an often-impossible standard to meet for work done in and between organizations. As with preregistration, encourage and support it when possible but do nothing to devalue research that is constrained from sharing data or fully disclosing proprietary methods and procedures. It is easy to imagine a research community's stealthy development of norms that "good" research is that which complies with tactics rooted in the principles of open science and, by implication, noncomplying research is less trustworthy. Vigilance against such a development is required because of the potential harm to scientific advancement in science-practice disciplines.

Summary and conclusion

This article is an attempt to emphasize approaches to research that encourage rigor, relevance, and progress. We share a critical value underlying open science, that of the necessary trustworthiness of scientific evidence and theory. Our essential concern is that implementing open science principles intended to realize that value would all too often undercut it and close doors to the conduct and publication of excellent research in science–practice domains. There are many reasons for this concern. They include, to name a few, discounting alternatives to hypothetico-deductive theorizing; imposing demands for disclosure and sharing that cannot be met; emphasizing a narrow approach to learned inquiry; failing to appreciate the evolving nature of big-data research in applied settings and the huge opportunity for the advancement of knowledge made possible by the proliferation of workforce, business, and customer/market data; undervaluing research that is practice-issue driven rather than theory driven; and creating a looming redefinition of the norms that define good research and theory. We are not alone in offering critiques of open science, many of which we have cited. In contrast to prior articles, we focus on potential hazards to the advancement of science in research-based, science–practice disciplines. Prior articles have largely ignored how the contributions to science of the practitioner–researcher who wishes to publish are being affected by open science practices. The recommendations we offer, if implemented, can help create a world in which applied research can substantially contribute to scientific advancement.

At least one major goal of the organizational sciences is to contribute to the development of evidence-based management (e.g., Rousseau, 2012). A science–practice foundation for that could develop along the lines of evidence-based medicine. Controlled experimental research, such as randomized control trials, provides an important evidentiary base for practice in medicine, as do large-sample epidemiological studies. Such epidemiological studies, in fact, share many attributes with big-data research in organizations: large numbers of participants, phenomena studied over time, numerous control variables, assessment of subgroup differences, and methods of analysis designed to provide strong support for cause-and-effect inferences. Additionally, case studies and clinical experiences often are the source of new discoveries and theories in the medical sciences. Indeed, frontline practitioners are the first to recognize fact patterns that signal the emergence of new diseases or useful new diagnostic categories. A further reality is that employers, particularly larger ones, preside over recurring opportunities in the form of naturally occurring experiments of workforce management and organization that provide deep insight into factors that influence the efficacy of specific practices or characteristics. Such opportunities can illuminate central propositions of the organizational sciences. Why would we want roadblocks to the ability of scientists to study and learn from these experiments? Our point is not to stretch an analogy between evidence-based medicine and management. We seek only to remind us all of the value of maintaining an openness to multiple approaches to discovery, testing, and theorizing in science–practice domains.

References

- Aguinis, H., Banks, G. C., Rogelberg, S. G., & Cascio, W. F. (2020). Actionable recommendations for narrowing the science–practice gap in open science. *Organizational Behavior and Human Decision Processes*, 158(May), 27–35. <https://doi.org/10.1016/j.obhdp.2020.02.007>
- Aguinis, H., Cummings, C., Ramani, R. S., & Cummings, T. G. (2020). “An A is an A”: The new bottom line for valuing academic research. *Academy of Management Perspectives*, 34(1), 135–154. <https://doi.org/10.5465/amp.2017.0193>
- Baker, M., & Penny, D. (2016). Is there a reproducibility crisis? *Nature*, 533(May), 452–454.
- Banks, G. C., Field, J. G., Oswald, F. L., O’Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 questions about open science practices. *Journal of Business Psychology*, 34(May), 257–270. <https://doi.org/10.1007/s10869-018-9547-8>
- Bauchner, H. (2018). Notice of retraction: Wansink B, Cheney MM. Super bowls: Serving bowl size and food consumption. *JAMA*. 2005;293(14):1727–1728. *Journal of the American Medical Association*, 320(16), 1648. <https://jamanetwork.com/journals/jama/fullarticle/2703449>

- Behfar, K., & Okhuysen, G. A.** (2018). Discovery within validation logic: Deliberately surfacing, complementing, and substituting abductive reasoning in hypothetico-deductive inquiry. *Organization Science*, *29*(2), 323–340.
- Berg, J.** (2017). Addendum to “Editorial retraction of the report ‘Environmentally relevant concentrations of microplastic particles influence larval fish ecology,’ by O. M. Lönnstedt and P. Eklöv.” *Science*, *358*(1630), 1549. <https://doi.org/10.1126/science.aar7766>
- Bray, D. W., Campbell, R. J., & Grant, D. L.** (1974). *Formative years in business: A long-term AT&T study of managerial lives*. Wiley.
- Campbell, J. P., & Wilmot, M. P.** (2018). The functioning of theory in industrial, work and organizational psychology (IWOP). In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work & organizational psychology: Personnel psychology and employee performance* (pp. 3–38). Sage.
- Carey, B.** (2015, August 27). Many psychology findings not as strong as claimed, study says. *New York Times*. <https://www.nytimes.com/2015/08/28/science/many-social-science-findings-not-as-strong-as-claimed-study-says.html>
- Cascio, W., & Aguinis, H.** (2008). Research in industrial and organizational psychology from 1963 to 2007: Changes, choices and trends. *Journal of Applied Psychology*, *93*(5), 1062–1081. <https://doi.org/10.1037/0021.9010.93.5.1062>
- Center for Open Science.** (2020a). *Participating Journals*. Retrieved March 13, 2020, from <https://cos.io/tr/>
- Center for Open Science.** (2020b). *Journals and Societies*. Retrieved March 13, 2020, from <https://cos.io/our-communities/journals-and-societies/>
- Chen, G.** (2015). Editorial. *Journal of Applied Psychology*, *100*(1), 1–4. <http://doi.org/10.1037/apl0000009>
- Chen, G.** (2018). Editorial: Supporting and enhancing scientific rigor. *Journal of Applied Psychology*, *103*(4), 359–361. <http://doi.org/10.1037/apl0000313>
- Cortina, J. M., Aguinis, H., & DeShon, R. P.** (2017). Twilight of dawn or of evening? A century of research methods in the Journal of Applied Psychology. *Journal of Applied Psychology*, *102*(3), 274–290. <http://doi.org/10.1037/apl0000163>
- Crandall, C. S., & Sherman, J. W.** (2015). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, *66*(September), 93–99. <http://doi.org/10.1016/j.jesp.2015.10.002>
- Dieterly, D., & Schneider, B.** (1974). The effects of organizational environment on perceived power and climate: A laboratory study. *Organizational Behavior and Human Performance*, *11*(3), 316–337. [https://doi.org/10.1016/0030-5073\(74\)90023-3](https://doi.org/10.1016/0030-5073(74)90023-3)
- Directory of Open Access Journals.** (2020). *Home page*. Retrieved March 13, 2020, from <https://doaj.org/>
- Douven, I.** (2017). Abduction. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2017 ed.). <https://plato.stanford.edu/archives/sum2017/entries/abduction/>
- Eden, D.** (2017). Field experiments in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*, 91–122. <https://doi.org/10.1146/annurev-orgpsych-041015-064400>
- Erland, B., Gross, S., & Guzzo, R. A.** (2017). Taco Bell enhances its people strategy with a new analytics recipe. Presentation at WorkatWork Total Rewards Conference & Exhibition, Washington, DC, May 9.
- Gabriel, A. S., & Wessel, J. L.** (2013). A step too far? Why publishing raw datasets may hinder data collection. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*(3), 287–290. <https://doi.org/10.1111/iops.12051>
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D.** (2016). Comment on “Estimating the reproducibility of psychological science.” *Science*, *351*(6277), 1037a–1037b.
- Göbel, C.** (2019). *Open citizen science—outlining challenges for doing and refining citizen science based on results from DITOs project*. Forum Citizen Science. doi:10.17605/osf.io/7etks
- Grand, J. A., Rogelberg, S. G., Allen, T. D., Landis, R. S., Reynolds, D. H., Scott, J. C., Tonidandel, S., & Truxillo, D. M.** (2018). A systems-based approach to fostering robust science in industrial-organizational psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *11*(1), 4–42. <https://doi.org/10.1017/iop.2017.55>
- Grand, J. A., Rogelberg, S. G., Banks, G. C., Landis, R. S., & Tonidandel, S.** (2018). From outcome to process focus: Fostering a more robust psychological science through registered reports and results-blind reviewing. *Perspectives on Psychological Science*, *13*(4), 448–456.
- Guzzo, R. A.** (2011). The universe of evidence-based I-O psychology is expanding. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *4*(1), 65–67. <https://doi.org/10.1111/j.1754-9434.2010.01298.x>
- Guzzo, R. A.** (2016). How big data matters. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 336–349). Routledge.
- Guzzo, R. A., Nalbantian, H. H., & Anderson, N.** (2022). Age, tenure, and business performance: A meta-analysis. *Work, Aging and Retirement*, *8*(2), 208–223. <https://doi.org/10.1093/workar/waab039>
- Guzzo, R. A., Nalbantian, H. N., & Parra, L. F.** (2014). A big data, say-do approach to climate and culture: A consulting perspective. In B. Schneider & K. Barbera (Eds.), *Oxford handbook of climate and culture* (pp. 197–211). Oxford University Press.
- Haig, B. D.** (2005). An abductive theory of scientific method. *Psychological Methods*, *10*(4), 371–388. <https://doi.org/10.1037/1082-989X.10.4.371>
- Hambrick, D. C.** (2007). The field of management’s devotion to theory: Too much of a good thing? *Academy of Management Journal*, *50*(6), 1346–1352. <https://doi.org/10.5465/AMJ.2007.28166119>

- Hausknecht, J. P., & Li, H.** (2018). Big data in turnover and retention. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 250–271). Routledge.
- Heilman, M. E., & Guzzo, R. A.** (1978). The perceived cause of work success as a mediator of sex discrimination in organizations. *Organizational Behavior and Human Performance*, *21*(3), 346–357. [https://doi.org/10.1016/0030-5073\(78\)90058-2](https://doi.org/10.1016/0030-5073(78)90058-2)
- Hernandez, I., Newman, D. A., & Jeon, G.** (2016). Methods for data management and a word count dictionary to measure city-level job satisfaction. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 64–114). Routledge.
- Hollenbeck, J. R., & Wright, P. M.** (2017). Harking, sharking, and tharking: Making the case for post hoc analysis of scientific data. *Journal of Management*, *43*(1), 5–18. <https://doi.org/10.1177/0149206316679487>
- Hong, Y., Liao, H., Hu, J., & Jiang, K.** (2013). Missing link in the service profit chain: A meta-analytic review of the antecedents, consequences, and moderators of service climate. *Journal of Applied Psychology*, *98*, 237–267. <https://doi.org/10.1037/a0031666>
- Illingworth, A. J., Lippstreu, M., & Deprez-Sims, A.-S.** (2016). Big data in talent selection and assessment. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 213–249). Routledge.
- Ionnidis, J. P. A., Stanley, T. D., & Doucouliagos, H.** (2017). The power of bias in econometrics research. *Economic Journal*, *127*(605), F236–F265. <https://doi.org/10.1111/ecco.jl2461>
- Kepes, S., & McDaniel, M. A.** (2013). How trustworthy is the scientific literature in industrial and organizational psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*(3), 252–268. <https://doi.org/10.1111/iops.12045>
- Kerr, N. L.** (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4
- King, E. B., Tonidandel, S., Cortina, J. M., & Fink, A. A.** (2016). Building understanding of the data science revolution and I-O psychology. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 1–15). Routledge.
- King, G., & Persily, N.** (2019). A new mode for industry-academic partnership. *PS: Political Science and Politics*, *53*(4), 703–709. <https://doi.org/10.1017/S10490965190001021>
- Klein, K. J., & Zedeck, S.** (2004). Introduction to the special issue on theoretical models and conceptual analysis. Theory in applied psychology: Lessons Learned. *Journal of Applied Psychology*, *89*(6), 931–933. <https://doi.org/10.1037/0021-9010.89.6.931>
- Kühberger, A., Fritz, A., & Scherndl, T.** (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS ONE*, *9*(9), Article e105825. doi:10.1371/journal.pone.0105825
- Leavitt, K.** (2013). Publication bias might make us untrustworthy, but the solutions may be worse. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*(3), 290–295. <https://doi.org/10.1111/iops.12052>
- Locke, E. A.** (2007). The case for inductive theory building. *Journal of Management*, *33*(6), 867–890. <https://doi.org/10.1177/01492063070636>
- Marszalek, J. M., Barber, C., Kohlhart, J., & Holmes, C. B.** (2011). Sample size in psychological research over the past 30 years. *Perceptual and Motor Skills*, *112*(2), 331–348. <https://doi.org/10.2466/03.11.PMS.112.2.331-348>
- McAbee, S. T., Landis, R. S., & Burke, M. I.** (2016). Inductive reasoning: The promise of big data. *Human Resource Management Review*, *27*(2), 277–290. <https://doi.org/10.1016/j.hrmr.2016.08.005>
- McShane, B. B., & Böckenholt, U.** (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research*, *43*(6), 1048–1063. <https://doi.org/10.1093/jcr/ucw085>
- Meyer, H. H., Kay, E., & French, J. R. P., Jr.** (1965). Split roles in performance appraisal. *Harvard Business Review*, *43*(January), 123–129.
- Murphy, K. R., & Aguinis, H.** (2019). HARKing: How badly can cherry-picking and trolling produce bias in published results? *Journal of Business and Psychology*, *34*(February), 1–19. <https://doi.org/10.1007/s10869-017-9524-7>
- Nalbantian, H., Guzzo, R. A., Kieffer, D., & Doherty, J.** (2004). *Play to your strengths: Managing your internal labor markets for lasting competitive advantage*. McGraw-Hill.
- Nalbantian, H. R., & Schotter, A.** (1995). Matching and efficiency in the baseball free-agent system: An experimental examination. *Journal of Labor Economics*, *13*(1), 1–31. <https://doi.org/10.1086/298366>
- Nalbantian, H. R., & Schotter, A.** (1997). Productivity under group incentives: An experimental study. *American Economic Review*, *87*(3), 314–341. <http://www.jstor.org/stable/2951348>
- Nalbantian, H. R., & Szostak, A.** (2004). How Fleet Bank fought employee flight. *Harvard Business Review*, *82*(4), 116–125. <https://hbr.org/2004/04/how-fleet-bank-fought-employee-flight>
- Nosek, B., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T.** (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Open Science Collaboration.** (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), Article aac4716. <https://doi.org/10.1126/science.aac4716>

- Oswald, F. L., Behrend, T. S., Putka, D. J., & Sinar, E. (2020). Big data in industrial-organizational psychology and human resource management: Forward progress for organizational research and practice. *Annual Review of Organizational Psychology and Organizational Behavior*, 7(January), 505–533. <https://doi.org/10.1146/annurev-orgpsych-032117-104553>
- Oswald, F. L., & Putka, D. J. (2016). Statistical methods for big data: A scenic tour. In S. Tonidandel, E. King, & J. Cortina (Eds.), *Big data at work: The data science revolution and organizational psychology* (pp. 43–63). Routledge.
- Pratt, M. G., Kaplan, S., & Whittington, R. (2020). Editorial essay: The tumult over transparency: Decoupling transparency from replication in establishing trustworthy qualitative research. *Administrative Science Quarterly*, 65(1), 1–19. <https://doi.org/10.1177/0001839219887663>
- Prosperi, M., Bian, J., Buchan, I. E., Koopman, J. S., Sperrin, M., & Wang, M. (2019). Raiders of the lost HARK: A reproducible inference framework for big data science. *Palgrave Communications*, 5(October), Article 125. <https://doi.org/10.1057/s41599-019-0340-8>
- Pulakos, E. D., Kantrowitz, T., & Schneider, B. (2019). What leads to organizational agility? ... It's not what you think. *Consulting Psychology Journal*, 71(4), 305–320. <https://doi.org/10.1037/cpb0000150>
- Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Harvard University, Graduate School of Business Administration.
- Rousseau, D. (Ed.). (2012). *The Oxford handbook of evidence-based management*. Oxford University Press.
- Schmidt, F. L. (2010). Detecting and correcting the lies that data tell. *Perspectives on Psychological Science*, 5(3), 233–242. <https://doi.org/10.1177/1745691610369339>
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62(5), 529–540. <https://doi.org/10.1037/0021-9010.62.5.529>
- Schneider, B. (1973). The perception of organizational climate: The customer's view. *Journal of Applied Psychology*, 57(3), 248–256. <https://doi.org/10.1037/h0034724>
- Schneider, B. (1975). Organizational climates: An essay. *Personnel Psychology*, 28(4), 447–479. <https://doi.org/10.1111/j.1744-6570.1975.tb01386.x>
- Schneider, B., Ashworth, S. D., Higgs, A. C., & Carr, L. (1996). Design, validity and use of strategically focused employee attitude surveys. *Personnel Psychology*, 49(3), 695–705. <https://doi.org/10.1111/j.1744-6570.1996.tb01591.x>
- Schneider, B., Hanges, P. J., Smith, D. B., & Salvaggio, A. N. (2003). Which comes first: Employee attitudes or organizational financial and market performance? *Journal of Applied Psychology*, 88(5), 836–851. <https://doi.org/10.1037/0021-9010.88.5.836>
- Schneider, B., Parkington, J.P., & Buxton, V.M. (1980). Employee and customer perceptions of service in banks. *Administrative Science Quarterly*, 25(2), 252–267. <https://doi.org/10.2307/2392454>
- Schneider, B., White, S. S., & Paul, M. C. (1998). Linking service climate and customer perceptions of service quality: Test of a causal model. *Journal of Applied Psychology*, 83(2), 150–163. <https://doi.org/10.1037/0021-9010.83.2.150>
- Schneider, B., Smith, D. B., Taylor, S., & Fleenor, J. (1998). Personality and organizations: A test of the homogeneity of personality hypothesis. *Journal of Applied Psychology*, 83(2), 462–470. <https://doi.org/10.1037/0021-9010.83.2.150>
- Shen, W., Kiger, W., Davies, T. B., Rasch, S. E., Simon, K. M., & Ones, D. S. (2011). Samples in applied psychology: Over a decade of research in review. *Journal of Applied Psychology*, 96(5), 1055–1064. <https://doi.org/10.1037/a0023322>
- Society for Industrial and Organizational Psychology. (2020). Website advertising. *The Industrial-Organizational Psychologist*, 58(2). <https://www.siop.org/Research-Publications/TIP/>
- Staw, B. M. (2016). Stumbling towards a social psychology of organizations: An autobiographical look at the direction of organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*, 3(March), 1–19. <https://doi.org/doi.org/10.annurev-orgpsych-041015-062524>
- Toth, A. A., Banks, G. C., Mellor, D., O'Boyle, E. H., Dickson, A., Davis, D. J., DeHaven, A., Bochantin, J., & Borns, J. (2020). Study preregistration: An evaluation of a method for transparent reporting. *Journal of Business and Psychology*, 36(June), 553–571. <https://doi.org/10.1007/s10869-020-09695-3>
- United Nations Educational, Scientific and Cultural Organization. (2022). *Global open access portal*. Accessed June 18, 2022. <https://www.unesco.org/en/natural-sciences/open-science>
- Vancouver, J. B. (2018). In defense of HARKing. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 11(1), 73–80. <https://doi.org/10.1017/iop.2017.89>
- Wade, N. (2012, January 11). University suspects fraud by a researcher who studied red wine. *New York Times*. <https://www.nytimes.com/2012/01/12/science/fraud-charges-for-dipak-k-das-a-university-of-connecticut-researcher.html>

Cite this article: Guzzo, RA., Schneider, B., and Nalbantian, HR. (2022). Open science, closed doors: The perils and potential of open science for research in practice. *Industrial and Organizational Psychology* 15, 495–515. <https://doi.org/10.1017/iop.2022.61>