Industrial and Organizational Psychology
PERSPECTIVES ON SCIENCE AND PRACTICE

A Step Too Far? Why Publishing Raw Datasets May Hinder Data Collection

ALLISON S. GABRIEL† AND JENNIFER L. WESSEL
The University of Akron

As junior faculty members, we understand the pressure to publish to ensure longevity in our academic careers. As Kepes and McDaniel (2013) articulate, this pressure may result in unethical practices, such as HARKing, suppressing findings, or, in extreme cases, utilizing data that is either fictitious or altered to support one's claims. Although these problems are real and not to be taken lightly, we believe that some of the suggestions made by Kepes and McDaniel may actually hinder our research efforts instead of help. In the discussion that follows, we outline our

Correspondence concerning this article should be addressed to Allison S. Gabriel.
E-mail: asg19@zips.uakron.edu
Address: Department of Psychology, The University of Akron, College of Arts and Sciences Building, Akron, OH 44325-4301
†Department of Management, Virginia Commonwealth University, 301 West Main Street, PO Box 844000, Richmond, VA, 23220

https://doi.org/10.1111/iops.12051 Published online by Cambridge University Press

concerns with the required submission and eventual publication of raw data as organizational scholars. Specifically, we believe that requiring raw data submission and publication will be problematic for three key reasons: (a) organizations may be less likely to ''buy in'' to projects, (b) sensitive populations may be less likely to consent to participating, and (c) responses may be distorted on sensitive variables. We then argue for increased accountability among coauthors to help minimize unethical publishing practices.

Potential Problems With Raw Data Submission/Publication

Organizational buy in. Top journals in the field often desire data coming from ''real'' employees as opposed to convenience samples of part-time employees to whom faculty in academia have easy access. As such, it becomes imperative to establish and strengthen ties with local organizations to obtain the elusive data that will generate top-tiered publications. From our experiences, this process entails a long courting phase where companies are given information about the study, provided the benefits of the study, and given information on how the company's identity, along with the employees' identities, will be kept anonymous. The goal is to let companies feel in control of not just the project (even if we are the ones collecting the data) but also aspects surrounding the data collection and data management. If our field were to shift toward a system in which all datasets would be submitted and eventually released to the public (even if they are deidentified datasets), this could present challenges in gaining company buy in.

Regardless of secure servers to protect the data for 5 years and anonymity of data collection, informing an organization that data are going to be collected *and* made available to other researchers (and potentially the ''public at large'') with whom they will never have contact will likely raise concern. This concern could result in rejecting project proposals to which they normally would have agreed, barring our field from the valuable information that could have been obtained. For instance, companies may be concerned about having their flaws exposed because the raw data could be reanalyzed by others or could have concerns about legal ramifications if employees were responding to sensitive material (an issue we return to in a subsequent section). Although requiring raw data submission and publication has the potential to stop unethical behavior, it also could prohibit data from being collected in the first place.

Sensitive populations. Even if organizations allow data collection, one must wonder what these data requirements would do to our ability to collect data from sensitive populations. As discussed in a recent focal article in this journal (Ruggs et al., 2013), there is a lot of work to be done in understanding the experiences of marginalized employees in the workplace. In attempting to recruit participants and gather data from these populations, it is imperative to ensure confidentiality (if not complete anonymity) due to the unique issues facing some marginalized groups. Specifically, marginalized populations with hidden identities (e.g., LGBT employees, employees with mental health diagnoses, employees with invisible physical illnesses, alcoholism, religious or political minorities) commonly fear the accidental disclosure of their identity at work and the potential negative consequences, from interpersonal discrimination by their coworkers to *legal* discrimination in certain areas of the world (Ragins, 2008). As such, studying employees with sensitive and/or hidden identities adds a significant barrier in the research process.

In our experiences of collecting data from these types of populations, there are unique challenges to (a) finding participants who are not readily visible to you and (b) convincing potential participants with sensitive identities that their data will remain secure and confidential. Informing participants from marginalized groups that their data may be publically available (even if identification information is removed)

https://doi.org/10.1111/iops.12051 Published online by Cambridge University Press

will add another barrier to collecting data from individuals who are already difficult to reach. Further, ensuring that data will be deidentified may not be enough to placate privacy concerns. In our opinion, additional barriers to obtaining data from these types of populations negatively affect our ability to fully understand the experiences of work for *all* types of employees and restricts our ability to examine structural and policy changes that could promote inclusion and fairness in the workplace.

Sensitive variables. Assuming that the company accepts the project and all participants willingly comply, a final concern would be assessing sensitive variables in employee surveys. In our own work, we have experienced situations where employees have concerns answering routine items such as turnover intentions (e.g., ''I already looked for another job,'' Sparr & Sonnentag, 2008), job satisfaction (e.g., ''In general, I like working here,'' Cammann, Fichman, Jenkins, & Klesh, 1983), and self-rated performance for fear of the information being linked back to them and communicated to upper management. In addition, individuals may be asked to report about their feelings of workplace incivility (from supervisors, coworkers, or customers) and discrimination at work. Moreover, in some studies focusing on constructs like performance appraisal and the feedback environment, employees may be asked to evaluate their supervisors' ability to provide feedback (e.g., ''My supervisor is often annoyed when I directly ask for performance feedback'') and level of supportiveness (e.g., ''When my supervisor gives me performance feedback, he/she is considerate of my feelings''),[1] which will inherently make employees concerned that their responses could ''get back'' to the boss. Similar to our concerns for marginalized populations, informing participants that raw data may be published may be a ''red flag'' for some who already question the intentions of the data collection.

1. Items are taken from the Feedback Environment Scale by Steelman, Levy, and Snell (2004).

Accountability Is Key Among Coauthors

It is hard to ignore the controversy surrounding Dierdrik Stapel and his colleagues; as referenced by Kepes and McDaniel, Stapel, unbeknownst to his collaborators, fraudulently made up datasets, many of which were published in top-tier, rigorous outlets (e.g., *Science*; *Psychological Science*). Although it may seem absurd that this type of data mistreatment could go unnoticed, it is highly plausible: As graduate students and faculty alike work on multiple projects, it is all too easy to divide the labor such that while some are writing others are analyzing. This inherently creates gaps in the research process and opens the door for unethical behaviors by some members of the research team. Hence, before we drastically alter the review process as Kepes and McDaniel suggest, we think it makes sense that enhancing accountability and ethics starts with us. As researchers, we need to abide by the same ethical code and uphold our colleagues to be equally accountable.

For instance, Beu and Buckley (2001) theorized that accountability likely enhances ethical behavior, given that individuals are put in a position where they must defend or explain their behavior to another person who has some kind of power and/or control over them. In the current context, being accountable to one's coauthors could enhance ethical behavior (i.e., utilizing data correctly) given that the coauthors (a) have reward power (i.e., could decide to remove a coauthor, could report a coauthor for unethical behavior) and (b) have shared expectations and/or responsibility for the publication and want to see that it is done correctly. As such, one of the first steps to helping ensure honest publishing is having all members of a publication verify that at least two members have knowledge of how the data were collected, checked the data management and clean-up process, and aided in the analyses for the manuscript. This could be done via a blanket statement that authors must sign upon submitting a manuscript. Although the ability to fake this part of the process would be present, part of

https://doi.org/10.1111/iops.12051 Published online by Cambridge University Press

the clause could be that authors, by signing and agreeing to the statement, acknowledge that any wrongdoing detected would result in equal penalties for all authors, as opposed to just a single perpetrator. This raises the incentives for all members to be honest because deceitful behaviors of one could be problematic for all.

Conclusion

Our way of publishing as scholars is at an ethical crossroads, with many viable options toward enhancing (and enforcing) ethical publishing. It is not our intent to dismiss suggestions made by Kepes and McDaniel. Rather, we encourage scholars to consider alternatives to our publication process with slight skepticism to make sure drastic changes do not lead to unintended obstacles to conducting research. In the short term, adjustments to coauthor accountability could help us travel down the more ethical academic road. Though no single suggestion nor single person, can recreate and improve our publication process, as a group, we have the ability to shift our collective moral compass to encourage proper publishing standards that hold us all accountable for high quality and impactful work.

References

Beu, D., & Buckley, M. R. (2001). The hypothesized relationship between accountability and ethical behavior. *Journal of Business Ethics*, *34*, 57–73.

Cammann, C., Fichman, M., Jenkins, G., & Klesh, J. (1983). Assessing the attitudes and perceptions of organizational members. In S. Seashore, E. Lawler, III, P. Mirvis, & C. Cammann (Eds.), *Assessing organizational change: A guide to methods, measures, and practices* (pp. 71–138). New York, NY: Wiley.

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in I–O psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*(3), 252–268.

Ragins, B. R. (2008). Disclosure disconnects: Antecedents and consequences of disclosing invisible stigmas across life domains. *Academy of Management Review Special Topic: Forum on Stigma and Stigmatization*, *33*, 194–215.

Ruggs, E. N., Law, C., Cox, C. B., Roehling, M. V., Wiener, R. L., Hebl, M. R., & Barron, L. (2013). Gone fishing: I–O psychologists' missed opportunities to understand marginalized employees' experiences with discrimination. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*, 39–60.

Sparr, J. L., & Sonnentag, S. (2008). Feedback environment and well-being at work: The mediating role of personal control and feelings of helplessness. *European Journal of Work and Organizational Psychology*, *17*, 388–412.

Steelman, L. A., Levy, P. E., & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement*, *64*, 165–184.

https://doi.org/10.1111/iops.12051 Published online by Cambridge University Press