# Generalizability challenges in applied psychological and organizational research and practice

**4 authors**, including:

Brenton M. Wiernik
University of South Florida
**88** PUBLICATIONS **1,284** CITATIONS

Mukhunth Raghavan
University of South Florida
**7** PUBLICATIONS **10** CITATIONS

**Some of the authors of this publication are also working on these related projects:**

Employee Contributions to Environmental Sustainability View project

psychmeta View project

# Generalizability Challenges in Applied Psychological and Organizational Research and Practice

Brenton M. Wiernik, Mukhunth Raghavan, Tyler Allan, and Alex J. Denison
*University of South Florida*

**Commentary on**
> Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences*.
> https://doi.org/10.1017/S0140525X20001685

**Abstract**
Yarkoni (2020) highlights patterns of overgeneralization in psychology research. In this comment, we note that such challenges also pertain to applied psychological and organizational research and practice. We use two examples—cross-cultural generalizability and implicit bias training—to illustrate common practices of overgeneralization from narrow research samples to broader operational populations. We conclude with recommendations for research and practice.

Yarkoni's (2020) critique focuses on overgeneralization from narrow sets of experimental stimuli, situations, or manipulations to very broad theoretical constructs. The issue of generalizability is not limited to academic and theoretical work but is also critically relevant for applied work that informs decisions made by organizations, institutions, and governments. Historically, applied psychology suffered from undergeneralization—a mistaken belief in "situational specificity," that effects were unique to specific contexts and never (or rarely) generalizable (Schmidt & Hunter, 1977). With the advent of meta-analysis, applied psychology learned that much of this apparent variability was due to statistical artifacts. However, the field may have overcorrected and now tends to overgeneralize. Frequently, models, measures, or interventions are "validated" in narrow settings, then applied in other contexts without carefully considering generalizability. Where generalizability is tested, it is often done in limited ways. In this commentary, use 2 key generalizability challenges as illustrations.

Brenton M. Wiernik wiernik@usf.edu ⓘ https://orcid.org/0000-0001-9560-6336, Mukhunth Raghavan mukhunth@usf.edu ⓘ https://orcid.org/0000-0002-4785-9387, Tyler Allan tylera1@usf.edu ⓘ https://orcid.org/0000-0002-9621-0461, Alexander J. Denison adenison@usf.edu ⓘ https://orcid.org/0000-0002-4291-8325

## Cross-Cultural Generalizability

First, consider cross-cultural generalizability. Much applied psychology research occurs in the U.S. and Western Europe, but these models are commonly used to motivate research and inform operational practices (e.g., selection systems, assessments, interventions) in organizations around the world (Gelfand et al., 2008). Researchers may allude to potential cross-cultural differences, but the core of Western models are generally assumed to apply across cultural contexts. Where generalizability across cultures is attended to, it is often done so haphazardly. A study might compare results in samples drawn from only two countries, rather than from a wider range of countries (Ones et al., 2012). Often, these countries are described as varying on a single dimension of cultural characteristics (e.g., collectivistic versus individualistic). Despite the narrow sampling of cultures in these studies, broad conclusions are often drawn about "collectivism" and assumed to apply to any culture that could be classified into these categories. Finally, samples in studies are frequently drawn from narrow subgroups within a culture (e.g., university students) without consideration of how these groups may differ from other cultural groups within a country. Several common overgeneralizations are apparent:

1. Individual countries are assumed to be exchangeable with others similarly classified.
2. The focal cultural characteristics (e.g., collectivism) are assumed to be the operative cause of cultural effects rather than other unmodeled factors (e.g., power distance, religiosity, history, economic environment).
3. The specific populations or subcultures sampled in a country are assumed to be representative of a country's cultural diversity.

These failures to consider included countries, characteristics, and subgroups as sampled from broader populations of these entities limit what conclusions are justifiable. If a researcher or organization observes that an intervention is effective in two contexts, they may conclude it will be effective more broadly. Conversely, if a study observes differences in relationships between two settings, they may *overestimate* variability in the relationship across cultures more broadly.

To better justify cross-cultural generalizations, researchers and practitioners must consider how representative their samples of individuals, characteristics, and countries are broadly. One possible approach is to is to conduct studies broadly sampling from diverse cultures around the world (Ones et al., 2012). By robustly sampling from many cultures, researchers can more accurately gauge whether relationships are consistent or variable across contexts. An alternative approach would be to extend conclusions about generalizability cautiously. Conclusions from narrow studies should be limited to the groups, countries, and contexts represented. In reports, investigators should conclude "X predicts Y in a sample of students in the United States" rather than making generalizations about broad cultural factors such as "collectivism". Over time, as single-context samples accumulate, systematic reviews of this evidence can identify the patterns that generalize (Oh, 2009; van Aarde et al., 2017). Such an approach encourages appropriate caution and also encourages de-centering of Western perspectives, allowing researchers themselves representing diverse cultures to pose questions that are relevant to their cultural contexts (Cheung et al., 2011; Gelfand et al., 2008).

## Organizational Interventions: The Case of Implicit Bias Training

Second, consider validation of interventions. Organizational interventions are frequently trialed in narrow populations (e.g., university students or employees from a few organizations), then deployed operationally without careful evaluation of their broader effectiveness. A recent popular example is implicit bias-based interventions to address issues of bias, racism, and inequities in organizations. Such interventions have become widespread, but evidence for their effectiveness for improving prejudice and inequity outcomes is sparse (FitzGerald et al., 2019; Onyeador et al., 2021). In a systematic review of implicit bias interventions, Forscher et al. (2019) found that the large majority of studies were conducted with U.S. university students, focused only on changes in implicit attitudes versus broader outcomes, and reported small effects. Importantly, Forscher et al. observed substantial heterogeneity across studies, underscoring that broad generalizability of implicit bias effects should not be expected. In light of this review, the uptake of implicit bias interventions for operational use has outpaced evidence supporting them. We identify three key areas of overgeneralization:

1. Generalization from studied populations (primarily university students) to operationally-relevant population (employees in specific industries and regions).
2. Generalization from studied outcomes (primarily short-term changes in implicit bias scores) to operationally-relevant outcomes (prejudice and equity outcomes).
3. Generalization from small observed effects to assume larger societal relevance. Despite observing small effects, studies frequently allude to potential large societal impacts (e.g., through accumulation across people or over time; cf. Oswald et al., 2015).

These overgeneralizations have consequences. Not only may these interventions be ineffective, but they appear to crowd out other actions that may better address systemic inequities (Pritlove et al., 2019). Organizational equity, diversity, and inclusion efforts should adapt to emphasize practices with stronger, more generalizable evidence bases such as intergroup contact and systems to bypass individual prejudice (Onyeador et al., 2021).

## Conclusion

We highlight these two examples as part of a broader pattern of overgeneralization in applied psychology from narrow samples, contexts, and measures to broader constructs and populations. To ensure effectiveness of organizational practices, we urge applied researchers and practitioners to generalizations more cautiously. In particular, we urge organizations to await evidence on operationally-relevant groups and measures (e.g., actual diversity and equity outcomes) before moving models and interventions into practice.

## References

Cheung, F. M., van de Vijver, F. J. R., & Leong, F. T. L. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, *66*(7), 593–603. https://doi.org/10.1037/a0022389

FitzGerald, C., Martin, A., Berner, D., & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: A systematic review. *BMC Psychology*, *7*(1), 29. https://doi.org/10.1186/s40359-019-0299-7

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. https://doi.org/10.1037/pspa0000160

Gelfand, M. J., Leslie, L. M., & Fehr, R. (2008). To prosper, organizational psychology should... Adopt a global perspective. *Journal of Organizational Behavior*, *29*(4), 493–517. https://doi.org/10.1002/job.530

Oh, I.-S. (2009). *The Five Factor Model of personality and job performance in East Asia: A cross-cultural validity generalization study* [Doctoral dissertation, University of Iowa]. http://search.proquest.com/dissertations/docview/304903943/

Ones, D. S., Dilchert, S., Deller, J., Albrecht, A.-G., Duehr, E. E., & Paulus, F. M. (2012). Cross-cultural generalization: Using meta-analysis to test hypotheses about cultural variability. In A. M. Ryan, F. T. L. Leong, & F. L. Oswald (Eds.), *Conducting multinational research projects in organizational psychology: Challenges and opportunities* (pp. 91–122). American Psychological Association. https://doi.org/10.1037/13743-004

Onyeador, I. N., Hudson, S. T. J., & Lewis, N. A. (2021). Moving beyond implicit bias training: Policy insights for increasing organizational diversity. *Policy Insights from the Behavioral and Brain Sciences*, *8*(1), 19–26. https://doi.org/10.1177/2372732220983840

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to predict ethnic and racial discrimination: Small effect sizes of unknown societal significance. *Journal of Personality and Social Psychology*, *108*(4), 562–571. https://doi.org/10.1037/pspa0000023

Pritlove, C., Juando-Prats, C., Ala-leppilampi, K., & Parsons, J. A. (2019). The good, the bad, and the ugly of implicit bias. *The Lancet*, *393*(10171), 502–504. https://doi.org/10.1016/S0140-6736(18)32267-0

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

van Aarde, N., Meiring, D., & Wiernik, B. M. (2017). The validity of the Big Five personality traits for job performance: Meta-analyses of South African studies. *International Journal of Selection and Assessment*, *25*(3), 223–239. https://doi.org/10.1111/ijsa.12175

Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences*. Advance online publication. https://doi.org/10.1017/S0140525X20001685