American Psychologist

ISSN: 0003-066X

2025, Vol. 80, No. 7, 1081–1085
https://doi.org/10.1037/amp0001569

COMMENTARY

Inconvenience and Generalization in Building a Better Psychology: Commentary on Sherman (2025)

Ivan Kroupin[1, 2], Joseph Henrich[2], and Michael Muthukrishna[1]
[1] Department of Psychological and Behavioural Science, London School of Economics and Political Science
[2] Department of Human Evolutionary Biology, Harvard University

In this commentary, we supplement Sherman's (2025) defense of convenience sampling, reviewing the complementary role of broad generalization and diverse samples. Specifically, Sherman's commentary could be misinterpreted as downplaying or disavowing the importance of broad generalization, despite the latter being necessary if we are to capture more than a narrow sliver of human cognitive variation. Moreover, we argue that stating the generalizability of our findings explicitly is key to both accurate interpretation and effective translation into applied work—a principle which holds even when our studies are not aiming to produce generalizable conclusions. We close with a review of practical ways in which broad generalization may be achieved. These include developmental, comparative, or computational approaches, as well as theoretical frameworks and "inconvenient" samples that capture cross-cultural variation.

Keywords: basic science; culture; convenience samples; Western, Educated, Industrialized, Rich, Democratic

Psychological science advances on two complementary fronts: discovery and generalization. Sherman (2025) rightly defends the value of convenience samples for the former. However, his article omits how generalization is necessary for psychological science as a whole and how considerations of generalization must shape our interpretation of empirical results—even those conducted without an aim to produce generalizable conclusions. In this commentary, we (a) affirm the importance of convenience samples as a tool, (b) discuss how generalization is key to psychological science as a whole and plays a role even in interpreting studies not intended to generalize across groups, and (c) illustrate the feasibility of broad generalization, including through the use of theoretical frameworks and "inconvenient" cross-cultural data sets.

Convenience Samples and Existence Proofs *Are* Necessary

Sherman (2025) reminds us that basic research often begins with an existence proof. When an effect is demonstrated "right here, right now," in a convenient sample, it can galvanize new theory. We concur. As Henrich et al. (2010) wrote,

> Exclusive use of WEIRD samples is justified when seeking existential proofs. … Our argument should not be construed to suggest that the exclusive use of WEIRD samples should always be avoided. There are cases where the exclusive use of these samples would be legitimate to the extent that generalizability is not a relevant goal of the research, at least initially. (Mook, 1983, p. 81)

Basic research need not claim universality from the outset, but we are concerned that Sherman (2025) might be interpreted

Rick Hoyle served as action editor.
Ivan Kroupin https://orcid.org/0000-0003-3673-9918
Joseph Henrich and Michael Muthukrishna contributed equally to this work. The authors have no known conflicts of interest to declare.

Ivan Kroupin played a leading role in conceptualization, writing–original draft, and writing–review and editing. Joseph Henrich played a supporting role in conceptualization and writing–review and editing. Michael Muthukrishna played a supporting role in conceptualization and writing–review and editing.
Correspondence concerning this article should be addressed to Ivan Kroupin, Department of Psychological and Behavioural Science, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom. Email: ivan@kroupin.com

as suggesting that convenience samples and existence proofs are sufficient even when broader generalization is desired. For instance, Sherman writes: "I advocate for reliance on convenience samples for basic psychological science" (p. 1074). Even though he notes that convenience samples are not always limited to Western, Educated, Industrialized, Rich, Democratic (WEIRD) populations, we argue that convenience samples alone can never be sufficient for generalization.[1] Limited sample diversity can offer insights only into (a) the specific population being tested and/or (b) those features of human psychology that are truly universal. Neither of these two outcomes alone provides a full understanding of mind and behavior (Kroupin et al., 2025). Moreover, convenience samples are typically insufficient to establish that a finding is universal: The psychological literature is rife with examples of apparently universal features of the mind that were later revealed to be culturally specific once researchers expanded beyond narrow samples (see Henrich et al., 2010; Kroupin, Davis, Zeng, & Henrich, 2024, for reviews).

In summary, if Sherman's (2025) argument is simply that convenience samples are acceptable when broad generalization is not the goal, then this supports the status quo: Researchers often begin with convenience studies and later broaden their samples. However, if his argument is read as implying that convenience samples suffice for basic research in psychological science, we must respectfully disagree. Given our inherently variable species, mapping the territory of the human mind necessarily involves capturing broad generalizations—both across all humans and across specific human groups. Doing so, in turn, ultimately demands evidence from beyond convenience samples, as we discuss in our final section.

Generalizability Matters (Even When We Are Not Generalizing)

As Sherman (2025) rightly points out, psychology has a long history of "sloppy language" (p. 1075)—that is drawing implicitly universal generalizations based on findings from WEIRD convenience samples (e.g., Cole, 1996; Henrich et al., 2010; Kroupin et al., 2025; Rad et al., 2018). Sherman addresses this issue by proposing that we clarify that the results of our studies (presumably those that do not aim to generalize beyond the sample) should be framed as simply "consistent with" a given theory, without implying their generalizability. This is certainly an improvement, but still fails to make explicit what scope of generalization we can draw from the findings. Making this scope explicit is not a question of formality or a quibble over phrasing—the only way we can understand where a given finding fits in the broader space of psychological science and thereby generate and test hypotheses about what other theories or mechanisms it should relate to, in which populations, and in which ways.

Moreover, appropriately stating generalizability is key to avoiding damaging mistakes in the translation of basic research into applied interventions. As Sherman (2025) correctly emphasizes, applied interventions have a crucial burden of proof when it comes to generalizability. Unfortunately, it is precisely the "sloppy language" of basic science which does not define the bounds of generalization, that often results in damaging generalizations by slipping into applied work unexamined. For instance, in the case of attachment theory, and early childhood development more generally, mechanisms studied in WEIRD societies and described in generalizing terms have been used to impose cultural models of childcare and parent–child attachment, which may be inappropriate outside of WEIRD contexts (Keller, 2018; Scheidecker et al., 2024).

To illustrate our point, consider two examples—(A) Sherman's (2025) study of stereotype use and cognitive load in U.S. undergraduates and (B) a study on attachment relations in a sample of U.S. mothers. Sherman-style conclusions for these two would read as follows:

A. "These data are consistent with *the theory that people use stereotypes more when they are under cognitive load*" (Sherman, 2025, p. 1075, emphasis in the original).

B. "These data are consistent with a theory that a single primary caregiver is important for normal development" (our formulation).

Contrast these with logically consistent formulations that make the scope of generalizability explicit:

A. "We have shown that there exist circumstances in which there is a relationship between cognitive load and stereotype use in U.S. undergraduates, consistent with a theory that this relationship holds generally in human cognition."

B. "We have shown that in a sample of U.S. mothers a single primary caregiver is important for normal development, consistent with a theory that such a single-primary relationship is important for development in all humans."

We would argue that the latter pair of conclusions does not diminish the importance of the results but does (a) more precisely relate the findings to the general theory that motivated the study (i.e., consistent with only a very local sample), (b) make clear that support for the general theory being tested would require empirical support for generalization (or details of how

[1] Note that if the role of convenience sampling is to support generalizations by combining with other convenience samples from diverse locations, then we are simply talking about diverse sampling spread across multiple studies.

the mechanism varies across groups), and (c) lessen the risk of inappropriate, potentially damaging generalizations leaking into applied work.

In sum, while we agree with Sherman (2025) that not all studies need to demonstrate generalizability, it is crucial to consider it when interpreting our findings (see also Simons et al., 2017, for a targeted call for adding constraints on generality to all academic papers). Doing so further highlights the need for studies that *do* attempt to generalize mechanisms broadly if we are to move toward a psychological science of humanity as a whole and not just a description of mechanisms in our convenience samples.

Is Broad Generalization Feasible?

Sherman (2025) acknowledges that we do not need to make truly universal claims for generalization to be an important goal. He also highlights the apparent implausibility of making any broad generalizations across groups given the massive cultural variation among humans. We agree that human psychological variation is vast, but we argue that the appropriate response is careful study, not despair (to paraphrase Cole et al., 1971). It is important to recognize that broad generalizability can never be *proven*—as with any hypothesis in any domain of science, we can only accumulate sufficient evidence to make a generalization plausible. Such evidence can come from at least three complementary approaches.

First, researchers can target cognitive systems that are likely to operate similarly across all humans by looking early in our ontogeny or phylogeny. For example, identifying a pattern in early infancy lends credence to the possibility that this pattern is present in all humans because there has been little time for postnatal environments to exert an influence (see, e.g., work on core knowledge, Carey, 2009; Spelke & Kinzler, 2007; though even here we must be cautious given evidence cultural differences can shape the brain even before birth, as in the case of language, Mariani et al., 2023). The case would be stronger still if the form of cognition is found across related species (see, e.g., Amir & Firestone, 2025, for an argument of this kind for certain visual illusions).

Second, we can develop formal theoretical models of how cognitive systems process information in general. Recent advances in Bayesian modeling and reinforcement learning frameworks, for instance, have proposed fundamental mechanisms by which cognitive systems may operate as a whole (e.g., Gershman, 2017; Silver et al., 2021; Tenenbaum et al., 2006). The coherence of the logic within these models, their complementarity to models of the evolution of these cognitive systems, and their predictive success all serve as evidence supporting the potential universality of the processes they describe.

These two approaches to generalizability are generally uncontroversial—though it is worth noting that they apparently contradict Sherman's (2025) claim that research supporting broad generalization is not attainable. Our position, however, is that developmental, comparative, and computational approaches, while necessary, are insufficient for the project of finding regularities within the space of human cognitive variation. After all, the approaches above are primarily useful in supporting *universal* generalization. However, it is abundantly clear that there are important generalizations at scales much smaller than all humans. Work on the nature of mechanisms underlying literacy and numeracy provides some of the clearest and most productive examples of research on such culturally specific features of the mind (Dehaene, 1997, 2010; see also Kroupin et al., 2025).

Thus, in addition to developmental, comparative, and modeling work, we argue that broad generalizations can and must be supported by generating theories of cross-cultural universality or variation, and by testing these theories empirically by collecting "inconvenient" samples from diverse populations. For instance, Henrich and Gil-White's (2001) theory of prestige predicts the existence of prestige-based status across human societies (Cheng et al., 2013; Henrich & Gil-White, 2001; see also von Rueden et al., 2011) and in other highly cultural species, including potentially in chimpanzees, elephants, and some toothed whales (Lee & Yamamoto, 2023).

Likewise, researchers who focused on abstract reasoning tasks have theorized that as institutions began to shape social life in ways that demanded (or permitted) greater abstract reasoning, educational institutions that enhanced this suite of cognitive abilities spread (e.g., Henrich, 2020; Muthukrishna & Henrich, 2016; see also Cole, 1996; Greenfield, 2009; Luria, 1976). Thus, if we are interested in making a claim about features of *human* abstract reasoning, then we would do well to test claims of human cognition across a spectrum of exposure to formal schooling. This would allow us to distinguish between cognition that reliably develops across all human societies and the cognition trained through the formal schooling system (see, e.g., Kroupin, Davis, Zeng, & Henrich, 2024; Rogoff, 1981; Scribner & Cole, 1973; Sharp et al., 1979, for relevant results and reviews).

This logic of theoretically structured sampling applies quite generally. For instance, we would do well to sample across the dimension of societies with tight versus loose norms (Gelfand et al., 2011) if we wish to make claims about human normative reactions, or across more rural versus urban environments for understanding variation in visual processing (e.g., Deręgowski, 2017; Kroupin, Davis, Lopes, et al., 2024). Returning to the example of attachment, sampling across groups with a diversity of childcare practices is key to an accurate picture of healthy child development across the globe (e.g., Keller, 2018, 2020; Lavelli et al., 2019). Heine and Norenzayan (2006) provided a classic, general discussion of theory and practice in systematic cross-cultural research, and our group has been engaged in a range of projects precisely with the goal of charting regularities in the space of human cognitive–

cultural variation (Atari et al., 2025; Atari & Henrich, 2023; Henrich, 2020; Kroupin & Zeng, 2024; Muthukrishna, 2023; Muthukrishna et al., 2021; Muthukrishna & Henrich, 2019).

Conclusion

We agree with Sherman (2025) that convenience samples continue to be a valuable tool in psychological science and that psychological science has often fallen short in how it makes and justifies broad generalizations. We also share Sherman's concern with providing a wider range of researchers with opportunities to pursue high-quality work. With this in mind, we have no objections to work with convenience samples where this is appropriate and cost-effective. Our suggestion, however, would be that working to improve the accessibility and funding for diverse sampling across a wider range of researchers is also key. After all, the diverse perspectives of diverse researchers may be particularly important in our efforts to identify and interpret variation across human cultural groups (see, e.g., Medin et al., 2010). In sum, our goal in this commentary has been to show how these issues interrelate with the need for generalization, as well as the rigorous sampling beyond convenient populations and the construction of theories that encompass and explain cultural variation (Muthukrishna & Henrich, 2019).

References

Amir, D., & Firestone, C. (2025). *Is visual perception WEIRD? The Müller-Lyer illusion and the cultural byproduct hypothesis*. PsyArXiv. https://doi.org/10.31234/osf.io/y7mtf

Atari, M., & Henrich, J. (2023). Historical psychology. *Current Directions in Psychological Science*, *32*(2), 176–183. https://doi.org/10.1177/09637214221149737

Atari, M., Henrich, J., & Schulz, J. (2025). *The chronospatial revolution in psychology*. PsyArXiv. https://doi.org/10.31234/osf.io/8atwz_v2

Carey, S. (2009). *The origin of concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. https://doi.org/10.1037/a0030398

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Harvard University Press.

Cole, M., Gay, J., Glick, J., & Sharp, D. W. (1971). *The cultural context of learning and thinking: An exploration in experimental anthropology*. Basic Books.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dehaene, S. (2010). *Reading in the brain: The new science of how we read* (Reprint ed.). Penguin Books.

Deręgowski, J. B. (2017). Cross-cultural studies of illusions. In A. G. Shapiro & D. Todorović (Eds.), *The Oxford compendium of visual illusions* (pp. 38–53). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199794607.003.0002

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Subirats Ferrer, M., Fischlmayr, I. C., … Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*(6033), 1100–1104. https://doi.org/10.1126/science.1197754

Gershman, S. J. (2017). On the blessing of abstraction. *Quarterly Journal of Experimental Psychology*, *70*(3), 361–365. https://doi.org/10.1080/17470218.2016.1159706

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, *45*(2), 401–418. https://doi.org/10.1037/a0014726

Heine, S. J., & Norenzayan, A. (2006). Toward a psychological science for a cultural species. *Perspectives on Psychological Science*, *1*(3), 251–269. https://doi.org/10.1111/j.1745-6916.2006.00015.x

Henrich, J. (2020). *The WEIRDest people in the world: How the west became psychologically peculiar and particularly prosperous*. Farrar, Straus and Giroux.

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196. https://doi.org/10.1016/S1090-5138(00)00071-4

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Keller, H. (2018). Universality claim of attachment theory: Children's socioemotional development across cultures. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11414–11419. https://doi.org/10.1073/pnas.1720325115

Keller, H. (2020). Children's socioemotional development across cultures. *Annual Review of Developmental Psychology*, *2*(1), 27–46. https://doi.org/10.1146/annurev-devpsych-033020-031552

Kroupin, I., Davis, H. E., & Henrich, J. (2025). Beyond Newton: Why assumptions of universality are critical to cognitive science, and how to finally move past them. *Psychological Review*, *132*(2), 291–310. https://doi.org/10.1037/rev0000480

Kroupin, I., Davis, H. E., Lopes, A. J. P., Konkle, T., & Muthukrishna, M. (2024). *Major urban–rural differences in perception of a visual illusion*. PsyArXiv. https://doi.org/10.31234/osf.io/2pu6f

Kroupin, I., Davis, H. E., Zeng, T. C., & Henrich, J. (2024). *Back to school: How formal education has shaped almost all of cognitive science*. PsyArXiv. https://doi.org/10.31234/osf.io/tqy2n

Kroupin, I., & Zeng, T. C. (2024). *From Richer to Leaner cultures and minds: How depth of standardization as an evolutionary dynamic drives changes in technology, social structure, and cross-cultural cognitive variation*. PsyArXiv. https://doi.org/10.31234/osf.io/mxfcv

Lavelli, M., Carra, C., Rossi, G., & Keller, H. (2019). Culture-specific development of early mother–infant emotional co-regulation: Italian, Cameroonian, and West African immigrant dyads. *Developmental Psychology*, *55*(9), 1850–1867. https://doi.org/10.1037/dev0000696

Lee, S. H., & Yamamoto, S. (2023). The evolution of prestige: Perspectives and hypotheses from comparative studies. *New Ideas in Psychology*, *68*, Article 100987. https://doi.org/10.1016/j.newideapsych.2022.100987

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Harvard University Press.

Mariani, B., Nicoletti, G., Barzon, G., Ortiz Barajas, M. C., Shukla, M., Guevara, R., Suweis, S. S., & Gervain, J. (2023). Prenatal experience with language shapes the brain. *Science Advances*, *9*(47), Article eadj3524. https://doi.org/10.1126/sciadv.adj3524

Medin, D., Bennis, W., & Chandler, M. (2010). Culture and the home-field disadvantage. *Perspectives on Psychological Science*, *5*(6), 708–713. https://doi.org/10.1177/1745691610388772

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, *38*(4), 379–387. https://doi.org/10.1037/0003-066X.38.4.379

Muthukrishna, M. (2023). *A theory of everyone: The new science of who we are, how we got here, and where we're going*. MIT Press.

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), Article 20150192. https://doi.org/10.1098/rstb.2015.0192

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

Muthukrishna, M., Henrich, J., & Slingerland, E. (2021). Psychology as a historical science. *Annual Review of Psychology*, *72*(1), 717–749. https://doi.org/10.1146/annurev-psych-082820-111436

Rad, M. S., Martingano, A. J., & Ginges, J. (2018). Toward a psychology of Homo sapiens: Making psychological science more representative of the human population. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11401–11405. https://doi.org/10.1073/pnas.1721165115

Rogoff, B. (1981). Schooling and the development of cognitive skills. In H. C. Triandis & A. Heron (Eds.), *Handbook of cross-cultural psychology* (Vol. 4, pp. 233–294). *Allyn & Bacon*.

Scheidecker, G., Tekola, B., Rasheed, M., Oppong, S., Mezzenzana, F., Keller, H., & Chaudhary, N. (2024). Ending epistemic exclusion: Toward a truly global science and practice of early childhood development. *The Lancet Child & Adolescent Health*, *8*(1), 3–5. https://doi.org/10.1016/S2352-4642(23)00292-4

Scribner, S., & Cole, M. (1973). Cognitive consequences of formal and informal education: New accommodations are needed between school-based learning and learning experiences of everyday life. *Science*, *182*(4112), 553–559. https://doi.org/10.1126/science.182.4112.553

Sharp, D., Cole, M., Lave, C., Ginsburg, H. P., Brown, A. L., & French, L. A. (1979). Education and cognitive development: The evidence from experimental research. *Monographs of the Society for Research in Child Development*, *44*(1–2), 1–112. https://doi.org/10.2307/3181586

Sherman, J. W. (2025). There is nothing WEIRD about basic research: The critical role of convenience samples in psychological science. *American Psychologist*, *80*(7), 1073–1080. https://doi.org/10.1037/amp0001443

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence*, *299*, Article 103535. https://doi.org/10.1016/j.artint.2021.103535

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*(6), 1123–1128. https://doi.org/10.1177/1745691617708630

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *278*(1715), 2223–2232. https://doi.org/10.1098/rspb.2010.2145

Received February 10, 2025
Revision received May 30, 2025
Accepted June 2, 2025 ■