

American Psychologist

Still WEIRD, Still Underreported: An Updated Benchmark for Psychological Science

Alison Jane Martingano, Neema Trivedi-Bateman, Lance Gauger, Megan Huggett, Beth Allcock, Rebecca E. Jackson, and Grace Holschuh

Online First Publication, April 6, 2026. <https://dx.doi.org/10.1037/amp0001703>

CITATION

Martingano, A. J., Trivedi-Bateman, N., Gauger, L., Huggett, M., Allcock, B., Jackson, R. E., & Holschuh, G. (2026). Still WEIRD, still underreported: An updated benchmark for psychological science. *American Psychologist*. Advance online publication. <https://dx.doi.org/10.1037/amp0001703>

Still WEIRD, Still Underreported: An Updated Benchmark for Psychological Science

Alison Jane Martingano¹, Neema Trivedi-Bateman², Lance Gauger¹, Megan Huggett², Beth Allcock²,
Rebecca E. Jackson^{2, 3}, and Grace Holschuh¹

¹ Department of Psychology, University of Wisconsin–Green Bay

² Department of Criminology, Sociology, and Social Policy, Loughborough University

³ School of Geography, University of Nottingham



Psychological science has long been critiqued for relying on samples drawn primarily from Western, educated, industrialized, rich, and democratic societies. The present study examines the current state of this critique by integrating three lines of evaluation: national sample origins, demographic representativeness, and adherence to American Psychological Association reporting standards. We analyzed all human subjects research published in 2024 in six leading psychology journals ($k = 1,217$) and compared their samples with previous reviews. While the proportion of U.S.-based samples continues to decline, U.S. participants remain vastly overrepresented (59% of samples in 2024). For demographic representativeness, we found substantial improvements in demographic reporting since 2014, but participants' race, ethnicity, socioeconomic status, and education are still frequently omitted. For reporting standards, we evaluated compliance with Journal Article Reporting Standards for Race, Ethnicity, and Culture (APA, 2023). We find wide variation across journals in the adoption of practices, with some, such as including sample justifications, almost unanimously adopted, whereas others, such as conducting analyses of diversity, are much less common. Overall, our findings show that psychology still depends heavily on Western, educated, industrialized, rich, and democratic samples, with only modest and uneven progress toward inclusivity and transparency. By combining three evaluation criteria, we establish a holistic benchmark that can be used to systematically monitor change in these research practices in the future.

Public Significance Statement

Psychological science has long grappled with ensuring that research samples accurately reflect the human population. Our study evaluates this issue along three strands: national origins of samples, demographic representativeness, and compliance with American Psychological Association Journal Article Reporting Standards for Race, Ethnicity, and

Joshua Clegg served as action editor.

Alison Jane Martingano  <https://orcid.org/0000-0001-8407-8611>

All data required to replicate analyses are available on the Open Science Framework at <https://osf.io/8cpb2/overview>. The authors have no known conflicts of interest to disclose.

Alison Jane Martingano played a lead role in conceptualization, formal analysis, methodology, project administration, software, visualization, writing—original draft, and writing—review and editing and an equal role in supervision. Neema Trivedi-Bateman played a supporting role in conceptualization, formal analysis, project administration, and writing—original draft and an equal role in supervision and writing—review and editing. Lance Gauger played a lead role in data curation, a supporting role

in methodology and writing—review and editing, and an equal role in writing—original draft. Megan Huggett played a supporting role in methodology and writing—review and editing and an equal role in data curation. Beth Allcock played a supporting role in methodology and writing—review and editing and an equal role in data curation. Rebecca E. Jackson played a supporting role in methodology and writing—review and editing and an equal role in data curation. Grace Holschuh played a supporting role in data curation and writing—review and editing.

 The data are available at <https://osf.io/8cpb2/overview>.

Correspondence concerning this article should be addressed to Alison Jane Martingano, Department of Psychology, University of Wisconsin–Green Bay, Mary Ann Cofrin Hall, Room C321, 2420 Nicolet Drive, Green Bay, WI 54311, United States. Email: martinga@uwgb.edu

Culture reporting guidelines, to provide a new holistic benchmark for evaluating the state of the field. The picture that emerges is sobering; U.S. samples still dominate; demographic reporting is still not universal; and adoption of Journal Article Reporting Standards for Race, Ethnicity, and Culture guidelines is uneven. Although there are areas for celebration, these findings reveal a misalignment between the field's stated commitment to diversity, rigor and transparency, and its actual publishing practices. Establishing this benchmark is a necessary first step toward holding the field accountable and enabling future researchers, journals, and policymakers to track meaningful change.

Keywords: diversity, sampling, representativeness, generalizability, open science

Supplemental materials: <https://doi.org/10.1037/amp0001703.supp>

Psychological science has faced sustained critiques of its reliance on narrow participant samples and limited reporting practices. Early evaluations documented the dominance of samples from Western, educated, industrialized, rich, and democratic (WEIRD) societies (e.g., Arnett, 2008), raising serious concerns about the generalizability of psychological research. Other scholars highlighted that even when research expanded beyond U.S. college students, samples were still often demographically unrepresentative (e.g., Rad et al., 2018). To compound the issue, many authors fail to report key demographic characteristics (e.g., race, socioeconomic status), making it difficult to assess sample representativeness (DeJesus et al., 2019; Rad et al., 2018). Recently, the American Psychological Association (APA) introduced new reporting standards for Journal Article Reporting Standards for Race, Ethnicity, and Culture (JARS-REC, American Psychological Association [APA], 2023), which formalize long-standing calls for fuller demographic reporting and ask authors to include sample justifications, consider constraints on generality, and adhere to transparency and openness among other recommendations. Taken together, these three threads, namely, sample national origins, demographic unrepresentativeness, and compliance with reporting standards, define three key challenges facing the field. This article brings them together to offer a holistic benchmark to evaluate the inclusivity and transparency of psychological research.

Unrepresentative National Origins

In the first major review of its kind, Arnett (2008) found that from 2003 to 2007, 68% of participants in top APA journals were from the United States, and 96% were from Western industrialized nations. U.S. samples were predominantly White and typically comprised of undergraduate students, often psychology majors (Arnett, 2008). Western university students were shown to be 4,000 times more likely to be studied than individuals outside the West (Henrich et al., 2010).

Henrich et al. (2010) noted that WEIRD participants differ from global populations in areas as diverse as visual perception, cooperation, and moral reasoning. These differences

suggest that WEIRD societies are among the least representative populations for generalizing about humans. More recent studies continue to affirm this, including wide-ranging research on antidepressant efficacy (Ruiz et al., 2023), freewill beliefs (Berniūnas et al., 2021), and algorithmic predictions (Li et al., 2021). These studies collectively reinforce the notion that, even in a globally interconnected and technologically advanced world, human psychological experiences are not homogeneous. Therefore, continued reliance on WEIRD samples raises concerns about the inappropriate generalization of psychological theories developed from narrow populations, potentially leading to misunderstandings of cognitive, emotional, and behavioral processes among diverse groups. Consequently, interventions and policies informed by such limited data may be less effective or even harmful when applied to populations whose experiences differ significantly from those typically studied.

The WEIRDness of human samples has been replicated across various subfields of psychology and social science, including developmental psychology (Nielsen et al., 2017), evolutionary psychology (Pollet & Saxton, 2019), clinical psychology (Kissel et al., 2024; Pedersen et al., 2022), and other social sciences such as political science, sociology, and economics (Jacobs & Mizrahi, 2020). Muthukrishna et al. (2020) even documented the prevalence of WEIRD samples in cross-cultural research, showing that even studies explicitly aiming to be cross-cultural still often disproportionately feature participants from WEIRD societies. Similarly, Krys et al. (2024) showed the tendency to draw conclusions about non-WEIRD cultures in general based on data from Confucian societies.

Following growing awareness of the issue, the decade after Arnett's (2008) review showed some little improvement in sample diversity. Rad et al. (2018), analyzing articles published in *Psychological Science* in 2014, found that 57.8% of samples were U.S. based, while the overall WEIRDness of samples remained largely unchanged (94.2%). Similarly, an analysis of eight high-impact social psychology journals comparing articles published between 1996–2000 and 2016–2020 found no significant change



Alison Jane Martingano

in the proportion of U.S. participants (67.2% vs. 64.3%; McGorray et al., 2023). In a broader cross-temporal effort, Thalmayer et al. (2021) revisited the same six journals examined by Arnett (2008) and reported that 62% of samples published between 2014 and 2018 were U.S. based, representing a modest improvement from the 68% observed a decade earlier.

Cross-temporal efforts are particularly valuable because they establish directly comparable time points using identical sampling frames. Retaining the same journal set minimizes variance introduced by differences in journal scope or impact and creates a clear opportunity to assess progress (or lack thereof) over time. To date, Arnett (2008) and Thalmayer et al. (2021) provided the most comprehensive and broadly applicable foundation for continued investigation, as both examine six flagship APA journals spanning multiple sub-fields of psychology. In selecting these journals as indicators of mainstream disciplinary practice, Arnett (2008) followed an established precedent in research on racial representation (Graham, 1992) and gender bias (Gannon et al., 1992), both of which similarly relied on premier APA journals to characterize the field. Arnett's (2008) review has since become a foundational critique of psychology's reliance on narrow participant populations, crystallizing concerns about the generalizability of psychological research. Continuing to document change using this same framework represents both a theoretically and methodologically coherent approach to tracking progress over time.

Demographic Representativeness

In terms of other demographic characteristics, the first decade following Arnett's (2008) review appears to have

brought little change as well. Rad et al. (2018) noted that the vast majority of articles included no demographic information other than gender. During this period, DeJesus et al. (2019) also found that most articles did not specify participants' race/ethnicity (73%), socioeconomic status (79%), or language background (74%). Some progress does appear to have been made in the most recent decade for the reporting of race and ethnicity (Bradford et al., 2024; McGorray et al., 2023; Medina et al., 2020), but not in other demographic features such as socioeconomic status (McGorray et al., 2023; Medina et al., 2020; Singh & Rajendra, 2024), and language (Medina et al., 2020).

Over the past 2 decades, there has been no shortage of calls for more representative samples. Notably, Bauer (2019, 2022) emphasized the need for systemic changes to include more diverse populations and improve reporting standards. Other authors have made excellent suggestions to achieve more representative samples, for example, advocating for stratified sampling and targeted recruitment to mitigate WEIRD biases (Muthukrishna et al., 2020; Protzko et al., 2026) and calling for policy changes in funding agencies to prioritize diverse sampling (Apicella et al., 2020). In addition, calls have been made for increased diversity of authors and editorial boards (Kahalon et al., 2021; Roberts et al., 2020; Singh et al., 2024; Solomonov et al., 2025), and there was hope that the future may bring increased representation (Forscher et al., 2022). However, whether the last decade has brought improvements remains unknown, and without regular, comparable analyses of participant diversity, it is difficult to know whether the field is moving in the right direction. Of note is the paucity of research in this area post-COVID-19, which shifted much research online (De Man et al., 2021) and therefore could have potentially increased diversity by transcending physical distance or further exacerbated the problem as samples would be limited to one side of the digital divide.

Compliance With Reporting Standards

Compounding the concerning realities of participant diversity, psychologists seldom offer cautionary notes about the source of their data or its potential cultural boundedness. In Arnett's (2008) original analysis, authors rarely mentioned the cultural context of the study, seemingly assuming that any human being could be taken to represent all people, a finding that was replicated a decade later by Rad et al. (2018). That same year, however, moves were made to correct this. The APA released updated Journal Article Reporting Standards (JARS) that required reporting of sample demographics (APA, 2018). More recently, the APA released the JARS-REC (APA, 2023). These standards were motivated by long-standing and well-documented concerns about the marginalization, misrepresentation, and erasure of people of



Neema Trivedi-Bateman

color in psychological science (APA, 2021). This includes deficit-oriented research questions, comparison norms that implicitly position White samples as the default, and uncritical extension of theories and measures developed in predominantly White, Western contexts to populations for whom they were neither designed nor validated (APA, 2021). In response to these concerns, JARS-REC were intentionally structured to target the stages of research and reporting at which problematic assumptions, inappropriate generalization, and measurement bias are most likely to be introduced or left unexamined. JARS-REC include 69 distinct reporting recommendations that span an entire article (title page through references). They address, among other issues, theoretical framing, sample justification, and adapting research instruments. JARS-REC were intended to function in addition to, rather than in place of, the earlier and broader JARS (APA, 2018). For example, while JARS require clear sample descriptions, JARS-REC specify how those descriptions should attend to culturally and nationally grounded definitions of race and ethnicity. Of course, whether authors have generally adopted these new recommendations is yet to be seen.

The Present Study

The present study builds on three strands of prior work. First, we provide an update on the national origins of psychological samples, extending previous reviews by Arnett (2008) and Thalmayer et al. (2021), by documenting samples in six top psychology journals in 2024. Second, we provide an update on the demographic characteristics of the samples, extending previous reviews (e.g., Rad et al., 2018). When possible, we compare 2024 data to these previous reviews. Finally, we move beyond these earlier reviews by assessing

whether authors are adopting recommended practices, most notably those outlined in JARS-REC (APA, 2023). By bringing these strands together, we provide a new benchmark against which future progress can be measured.

Method

Transparency and Openness

All data required to replicate analyses are available at <https://osf.io/8cpb2/overview>. This research was not pre-registered as our analyses were truly exploratory.

Literature Review of Six Core Psychology Journals

This study screened all articles published in 2024 in six top APA journals: *Developmental Psychology*, *Journal of Personality and Social Psychology*, *Journal of Psychopathology and Clinical Science* (formerly the *Journal of Abnormal Psychology*), *Journal of Family Psychology*, *Health Psychology*, and *Journal of Educational Psychology*. These journals were chosen to replicate Arnett's (2008) and Thalmayer et al.'s (2021) studies to allow comparison to prior work.

We excluded commentaries, review articles, and studies involving nonhuman subjects from our analysis, leaving a total of 551 original research articles. In Arnett's (2008) original coding scheme, articles that included samples from more than one region were coded separately for each region represented. When an article reported multiple studies conducted within the same region, however, it was coded only once for that country rather than separately for each study. To preserve comparability with this prior work, we adopted this same approach for our comparative analyses. However, given the increasing prevalence of multistudy articles, this convention may underrepresent regions that are examined repeatedly within a single publication. Accordingly, we also report k based on all samples in their entirety and use this full-sample k for all other analyses. To determine k , if an article included multiple studies, each study was coded separately. Samples were also coded separately if the authors presented the results for distinct groups independently. For example, if a study included both men and women, this was coded as a single sample unless the primary results were reported separately for men and women, in which case it would be treated as two samples. In the case of longitudinal studies that followed the same sample over time, only the details of the Time 1 sample were coded.

This approach yielded a total of 1,220 unique samples. Three samples were excluded, which contained over a quarter of a million participants each (from data sources such as a U.S. Department of Labor), which left 1,217 samples for analysis.



Lance Gauger

Article Coding

We coded articles across three domains: (a) national origin of samples, (b) demographics of samples, and (c) reporting standards of article. Coding was conducted in Covidence (Veritas Health Innovation, n.d.), an online platform designed to support systematic reviews by providing a structured environment for screening, data extraction, and resolution of coding disagreements. To standardize the process, we developed a structured extraction template, which all coders used, listing each variable to be coded alongside detailed coding instructions.

Coding was carried out by four trained undergraduate research assistants. Training was led by the first two authors through multiple recorded sessions with ongoing supervision. Articles were distributed across coder pairs so that articles were independently coded by two coders. Coder agreement ranged between 80.4% and 98.2%. Cohen's κ ranged from .03 to .63, in lower values reflecting extreme category imbalance and the rarity of some codes.

National Origin of Sample

The geographic location of each sample was coded using the same coding technique as those of Arnett (2008), Thalmayer et al. (2021), and Rad et al. (2018) so that our results were comparable: Europe, Asia, Latin America, Africa, the United States, English-speaking countries, Israel, and the Middle East. The United States was coded as a separate category to evaluate whether U.S.-based samples still dominated psychological research. The category of English-speaking countries was developed by Arnett (2008) to represent a group of countries with strong cultural and historical ties to the United States: the United Kingdom,

Canada, Australia, and New Zealand. Samples with unknown national location were coded as “unknown.”

Demographics of Sample

Sample demographics for each article were recorded: total sample size, the specific country or countries from which the sample originated, the average age of participants, gender representation, racial composition, socioeconomic status, education, language, and whether the data were collected online or offline. We also noted whether the sample included any special populations, such as individuals with a specific diagnosis or those in a particular profession.

Compliance With Reporting Standards

To minimize redundancy, we evaluate coded items in only one section. When JARS-REC items duplicate content assessed elsewhere, we report it in the first relevant content area only. For example, JARS-REC's directive to “describe the population(s) from which the sample was drawn” is covered under national origin of samples, and recommendations to report demographic characteristics are covered under demographics of samples. This section therefore focuses on JARS-REC elements not already captured above. JARS-REC includes nearly 70 distinct recommendations, many of which are highly overlapping; therefore, we grouped related items into broader coding categories that could be applied consistently across all articles. Please see Supplemental Materials for more details.

Sample Described in Abstract. This category reflected the JARS-REC recommendation that abstracts should acknowledge the racial and ethnic heterogeneity of the sample, ideally in combination with other pertinent characteristics relevant to the study. We coded each abstract according to the number of distinct demographics features it reported: one, two, or three or more features.

Sample Justification. This category combined multiple JARS-REC items concerning how authors explained and contextualized their sampling decisions. Each article was then coded as providing no justification, a theoretical justification (e.g., linking the choice of sample to a research question or framework), a practical justification (e.g., describing recruitment procedures, constraints, or oversampling strategies), or both when elements of each were present.

Appropriateness of Measures. This category combined multiple JARS-REC recommendations concerning whether the measures used were culturally and linguistically appropriate for the study sample. We operationalized this code conservatively: unless there was clear evidence that a measure was used inappropriately, we assumed appropriateness. For example, we coded measures as not appropriate if they were administered without translation to participants who did not speak the language or if instruments designed for adults were used with child samples.



Megan Huggett

Discussion of Constraints on Generality. This category combined multiple JARS-REC recommendations relating to how authors acknowledged the limits of generalizability and contextualized their findings. Each article was coded as providing either (a) no mention of constraints, (b) pro forma acknowledgment, or (c) thoughtful discussion. We coded as pro forma those statements that offered only a cursory acknowledgment of limitations without elaboration.

Analyses of Diversity. This category combined multiple JARS-REC recommendations relating to whether authors examined differences across demographic groups and interpreted those differences appropriately. Each article was coded dichotomously as yes (any attempt was made to analyze results by demographic subgroup or assess inequities) or no (no such analyses reported). Although we recognize that subgroup analyses should not be conducted when underpowered, we do not report on whether analyses were sufficiently powered. Determining the adequacy of statistical power would have required information that was not available in the published articles and thus could not be assessed reliably.

Transparency. This category reflected the JARS-REC recommendation that authors include a “Transparency and Openness” subsection detailing their use of preregistration, data repositories, and open materials. Because many articles did not have such a dedicated subsection, we coded the presence of these practices in any part of the article. Specifically, we recorded whether data were made public, whether materials were made public, and whether the study was preregistered.

We did not code all JARS-REC recommendations (see Supplemental Materials for more details). Several items could not be assessed reliably given the information available in published articles. JARS-REC guidelines are intended primarily for authors, who possess intimate knowledge of

their study design and intentions; external coders cannot always determine whether a recommendation is applicable or has been met. We therefore excluded recommendations in three main categories: (a) conditional (“If X, then Y”) items. Many JARS-REC recommendations are framed contingently (e.g., “highlight relevant constructs regarding race, ethnicity, and culture when they are the focus of the research questions under investigation.”). Because we did not know whether these conditions were relevant for a given study, we could not reliably code the “then” component. (b) Items requiring subject-matter expertise: Some recommendations involve judgments that cannot be made without insider knowledge or specialized disciplinary expertise. (e.g., “critique existing evidence on the generalizability . . . , alternatively, acknowledge the lack of research”). Determining whether there is or is not relevant research requires field-specific expertise and cannot be gleaned from the article alone. (c) Overly specific or niche items: We also excluded recommendations that would apply to only a very small fraction of studies and thus could not be meaningfully tracked at scale. For example, “avoid rote implementation of White participants as the reference group” is highly important but only relevant in studies that conduct specific cross-racial analyses. Including such items would have added noise without substantially advancing our goal of establishing a broadly applicable benchmark.

Extended Literature Review Across the Wider Field

To place our findings in context, we conducted a structured review of published evaluations of sample representativeness across journals. Using Arnett’s (2008) seminal article as an anchor point, we retrieved all 1,845 publications citing this work. After title screening, 110 articles remained, and following full-text review, 21 articles were identified that provided quantitative data on the nationality of samples (see Supplemental Materials for PRISMA-style diagram). Some reviews collapsed across multiple journals or did not specify the outlets in which studies were published. After excluding these, we obtained usable data from 12 articles. Because many articles reported results for multiple journals and/or multiple time points, this yielded 60 estimates of U.S. sample proportions and 62 estimates of WEIRD sample proportions. When combined with the current analyses, the final analytic samples included 62 journal-level estimates of U.S. samples and 68 of WEIRD samples. CiteScore data and the geographical location of journals were then retrieved for the relevant year of analysis. CiteScore is a journal metric that reflects the average number of citations received per article over a 4-year window. We used CiteScore as a standardized, widely available proxy for journal citation impact. This corpus covered a wide range of journals covering a variety of subfields within psychology with CiteScores ranging from 1.4 to 22.4 and from journals based in a variety of countries such as the United Kingdom, Belgium, and the Czech



Beth Allcock

Republic. However, no journals were from non-WEIRD countries. These data allowed us to pair the sample estimate for that journal (percent U.S. and/or WEIRD) with the journal's CiteScore from that same time point, allowing us to determine whether reliance on U.S. or WEIRD samples varies systematically across journals as a function of journal visibility, and ultimately infer if the patterns observed in APA journals are anomalous or reflect broader structural features of psychological publishing.

Results

Results are organized to first establish descriptive benchmarks for 2024, followed by comparative analyses with prior reviews when identical variables and coding schemes were available. Comparisons are restricted to cases wherein methodological alignment permits meaningful inference about change over time (e.g., national origin comparisons

with Arnett, 2008, and Thalmayer et al., 2021; demographic reporting comparisons with Rad et al., 2018). Variables not assessed in prior work are presented descriptively for 2024 only.

National Origin of Samples

National Origin in 2024

The majority of samples published in 2024 were recruited from the United States (59%; see Table 1), with notably large percentages from Europe (15%) and other English-speaking countries (12%). Sample sizes from these regions were also large, representing 50% (the United States), 15% (Europe), and 12% (English speaking) of participants sampled. When population size is taken into account, these disparities are magnified. Sampling density in the United States reached 295 participants per 100,000 residents, compared with rates of three per 100,000 in Asia and below one per 100,000 in Africa and the Middle East (see Table 1).

Comparison of Changes in National Origin Over 2 Decades

To assess whether sample distributions have changed over time, we conducted comparative analyses with two prior time points using identical coding schemes (Arnett, 2008; Thalmayer et al., 2021; see Table 2). Three chi-square tests of independence were conducted to examine whether the regional distribution of samples changed significantly over time. Results indicated statistically significant shifts in representation between 2003–2007 and 2014–2018, $\chi^2(7) = 40.24$, $p < .001$, and between 2014–2018 and 2024, $\chi^2(7) = 36.47$, $p < .001$, and across the entire period from 2003–2007 to 2024, $\chi^2(14) = 94.82$, $p < .001$. To determine which regions contributed most to these overall changes, standardized residuals were examined. Representation of U.S. samples has not changed significantly over either decade. European samples

Table 1
National Origin of Samples in 2024

Region	Sample <i>k</i> (%)	Participant <i>N</i> (%)	~Population in millions	<i>N</i> per 100,000
United States	714 (59)	1,004,103 (50)	341	295
Europe	181 (15)	296,748 (15)	750	40
Asia	107 (9)	124,306 (6)	4,800	3
Latin America	18 (2)	80,103 (4)	670	12
Africa	11 (1)	3,716	1,600	0.2
Middle East	5	2,535	452	0.6
English speaking	151 (12)	440,435 (22)	133	331
Israel	29 (2)	44,056 (2)	10	441
Unknown	51 (4)	18,693 (1)		

Note. Sample *k* = number of samples collected from that region. Participant *N* = number of participants included in the samples from that region. Percentages indicate the weight of that region as a percentage of all the regions. *N* per 100,000 indicates how many participants were recruited for every 100,000 individuals estimated to live in that region. English-speaking countries are the United Kingdom, Canada, Australia, and New Zealand.



Rebecca E. Jackson

showed increased representation in the first decade, followed by no change thereafter. In the latter decade, Asian samples have increased in representation (see Supplemental Materials).

We then examined changes in representation across individual journals. Five of the six journals demonstrated improvements in non-U.S. representation between 2003–2007 and 2024, apart from the *Journal of Psychopathology and Clinical Science* (see Supplemental Materials).

National Origin Across Journals

We examined whether journal impact was associated with the national origin of samples in that journal. Specifically, we correlated journals' CiteScore with the proportion of samples drawn from (a) the United States and (b) WEIRD regions. Bivariate correlations with CiteScore revealed no significant relationship with U.S. sample proportion, $r(60) = -.19, p = .156$, nor with WEIRD sample proportion, $r(66) = -.02, p = .862$ (see Figure 1 for scatterplot). This indicates that reliance on WEIRD

and U.S.-based samples is a systemic feature of psychological publishing, not concentrated in lower or higher impact outlets.

Independent-samples t tests comparing U.S.-based journals to non-U.S.-based journals revealed that WEIRD participants were no more likely to be included in U.S.-based than in non-U.S.-based journals, $t(64) = -0.75, p = .454$. However, U.S.-based journals were more likely to publish U.S. samples, $t(58) = -2.37, p = .021$. This suggests that the problem of WEIRD samples is not unique to U.S.-based journals; however, the reliance of U.S. samples specifically appears exaggerated in U.S.-based journals.

Demographics of Samples

Demographics of Samples in 2024

It is difficult to make conclusions about sample demographics beyond age and gender, as many participants' characteristics were not reported (see Figure 2). Gender was reported for 95.6% of samples, and age was reported in 91.4%. On average 41.7% of participants were men, 57.6% were women, and 0.5% reported another gender. The mean age of participants across studies was 26.04 years ($SD = 15.63$), with ages ranging from 0.01 to 80.20 years.

Language was reported for 23.7% of samples, with 11.8% reporting English, 1.9% German, 1.6% French, 1.0% Spanish, 0.8% Hebrew, and 4.1% reporting Bilingual samples. Despite the predominance of English, it may still be reasonable to assume this is an underrepresentation of the dominant language as many articles not reporting what language participants spoke could be reasonably inferred to have English-speaking participants. Race was reported for 50.53% of samples, with 61.72% of participants reporting being White. Socioeconomic status was reported even less frequently, with only 33.6% of studies providing this information. Education level was reported in 43.1% of cases.

Overall, 47.3% of samples were collected online. Many samples were comprised of "special groups," including children (19.6%), undergraduates (13.0%), and parents (11.8%). Other

Table 2

Comparative Analyses of National Origins of Samples Over 2 Decades as a Simple Count (N) and Including All Samples (k)

Region	2003–7 ^a	2014–18 ^b		2024	
	N (%)	N (%)	k (%)	N (%)	k (%)
United States	1,677 (68)	2,232 (63)	4,172 (66)	356 (57)	714 (59)
Europe	311 (13)	594 (17)	967 (15)	108 (17)	181 (15)
Asia	78 (3)	141 (4)	261 (4)	54 (9)	107 (9)
Latin America	13 (1)	25 (1)	30	9 (1)	18 (1)
Africa	5	25 (1)	27	8 (1)	11 (1)
Middle East	2	17	23	2	5
English speaking	335 (14)	488 (14)	724 (11)	78 (12)	151 (12)
Israel	31 (1)	45 (1)	97 (2)	10 (2)	29 (2)

Note. English-speaking countries are the United Kingdom, Canada, Australia, and New Zealand.

^a Arnett (2008). ^b Thalmayer et al. (2021).



Grace Holschuh

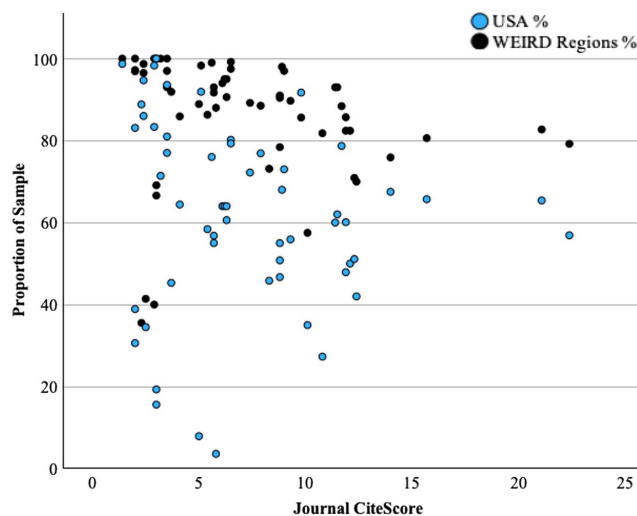
special groups included adolescents (4.9%), clinical populations (3.9%), and infants (3.4%).

Comparison of Changes in Demographics Over a Decade

To assess whether reporting practices have changed over time, we conducted comparative analyses for demographic variables that were evaluated in a prior review (Rad et al., 2018). For each demographic variable, a Fisher's exact test

Figure 1

The Relationship Between Journal Impact (CiteScore) and the Proportion of Samples Drawn From the U.S. (Blue) or WEIRD (Black) Regions



Note. Points represent journal-by-year estimates. WEIRD = Western, educated, industrialized, rich, and democratic. See the online article for the color version of this figure.

was conducted to examine whether the proportion of studies reporting this information had improved. The results showed a significant increase in the reporting of gender (from 75% to 95.5%, $OR = 7.05$, 95% CI [5.32, 9.33], $p < .001$), race/ethnicity (from 20% to 51.9%, $OR = 4.32$, 95% CI [3.76, 4.97], $p < .001$), and socioeconomic status (from 8% to 33%, $OR = 5.68$, 95% CI [4.79, 6.73], $p < .001$). Reporting of education decreased (from 54% to 43%, $OR = 0.64$, 95% CI [0.56, 0.73], $p < .001$).

Compliance With Reporting Standards

Compliance With Reporting Standards in 2024

The results reported here are complementary to the results in earlier sections, which simultaneously captured compliance with JARS-REC standards on population and demographics reporting.

Most research included some information about the sample in the abstract of the publication, with only 9.6% having no information at all. Only sample size was reported in the abstract by 13.4% of articles with 76.4% reporting additional demographic features (17.8% one feature, 23.6% two features, 34.9% three features). The reporting of demographics in the abstract differed by journal (see Table 3). Almost all research (97.6%) provided either a theoretical or practical justification for their sample choice.

Almost all studies (99.6%) met the minimal criteria for measure appropriateness (e.g., translated, age appropriate), although in most cases this was because the standard measure was designed for and used with a WEIRD sample.

Most research acknowledged the constraints on generality relevant to the sample they used in their article (86.7%) with 50.0% providing a detailed and thoughtful examination of the groups to which the results may or may not be relevant and 36.0% giving some (often pro forma) acknowledgment.

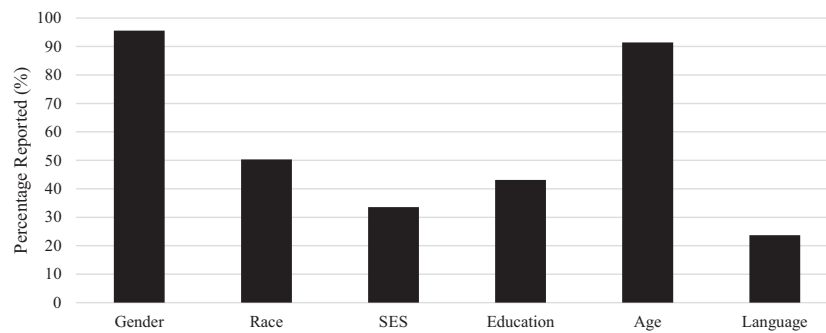
Some analyses of diversity among their sample were conducted by 60.4% of articles, for example, examining if their main finding differed by gender.

Transparency was most common in the form of publicly available data (50.3%), followed by publicly available materials (44.7%). Preregistration differed dramatically by journal, ranging from 12.9% in *Journal of Family Psychology* to 73.9% in *Journal of Personality and Social Psychology*, with an overall average across all six journals of 29.1% (see Table 3).

Discussion

Psychological science remains WEIRD in terms of both sample origins and participant demographics, continues to underreport those features, and is characterized by inconsistency in adopting best practices for reporting standards. However, there is some room for (modest) celebration. Across six leading journals in 2024, U.S. and other WEIRD

Figure 2
Proportion of Samples Reporting Key Demographic Variables in 2024



Note. SES = socioeconomic status.

samples continue to dominate, but a slow movement toward diversification is observable. However, demographic diversity remains hard to evaluate beyond age and gender as most studies continue to omit other demographic characteristics. Moreover, despite the introduction of JARS-REC guidelines designed to encourage fuller contextualization and transparency, the adoption of these practices is uneven across journals. There has been commendable adoption of some recommendations, with samples being frequently justified in text and at least some constraints on generality acknowledged. Taken together, these findings establish a relatively low benchmark of current practice, highlighting substantial room for improvement in both sample diversity and reporting standards.

National Origin of Samples

We observed a slow shift in the national origin of research samples over the past 2 decades, toward greater diversity, with the overrepresentation of U.S. samples decreasing by roughly 10% points. When adjusting for population size, U.S. participants are still 25 times more likely to be participants in psychological research than people living elsewhere. This stark overrepresentation highlights the continued reliance on WEIRD populations, raising concerns about the generalizability of research findings to the global population. Representation, however, has not proceeded universally with most improvement occurring because of increasing the representation of Asian samples (from 3% to 9%). At the same time, regions like Africa, Latin America, and the Middle East showed little to no change. These results are in line with cultural research, which has noted that cross-cultural studies frequently fail to adequately represent non-Confucian societies (Krys et al., 2024).

Overall, it appears that efforts to diversify the national origins of participants in psychological research have yielded only sparse results. This only serves to reinforce previous commentary that “American psychology is becoming more international but too slowly” (Webster et al., 2021, p. 1).

Demographics of Samples

Another key finding in this study is the ongoing inadequacy in the reporting of demographic characteristics beyond basic variables like gender and age. In particular, race/ethnicity, socioeconomic status, and educational background remain underreported. Race/ethnicity was reported in 50.5% of samples and socioeconomic status in only 33.6%. Although these numbers reflect a substantial improvement in reporting over the last decade (Rad et al., 2018) and even from 5 years ago (DeJesus et al., 2019), the level of reporting is still concerning, especially given the importance of these variables for understanding the generalizability of psychological findings across diverse populations.

It is important to note that it is not easy to collect race/ethnicity data in specific settings, such as in France, the Netherlands, or Germany, due to legal restrictions (Singh et al., 2024). Nevertheless, it is important to use contextually appropriate questions for demographic data collection. For example, in China, asking participants to report ethnicity instead of race as the country is largely monoracial but has upward of 56 distinct ethnicities. Other recommendations include collecting additional sample descriptions relevant to theoretical questions and sharing sociodemographic data in ways that are consistent with de-identification standards (Singh et al., 2024).

Compliance With Reporting Standards

APA released updated JARS in 2018 after a decade of calls for improved methodological practices in psychological science spurred by the lack of reproducibility (Open Science Collaboration, 2015). More specifically, JARS-REC were introduced in 2023 to address persistent shortcomings in demographic reporting and to encourage researchers, among other things, to justify sampling decisions, use culturally appropriate measures, and acknowledge the inevitable constraints on generalizability. In this way, JARS-REC serves as the connective tissue across our three strands, recommending

Table 3
Methodology and Reporting Among Six Journals in Percentages

Journal	Demographics in abstract	Sample justification	Appropriate measure	Constraint acknowledged	Analysis of diversity	Public data	Public material	Preregistered
<i>Developmental Psychology</i>	99.5	100.0	100.0	87.0	66.4	49.3	41.8	19.2
<i>Health Psychology</i>	97.5	97.5	100.0	87.5	76.3	38.8	46.3	30.0
<i>Journal of Educational Psychology</i>	88.7	96.2	98.7	80.0	47.5	61.3	50	16.3
<i>Journal of Family Psychology</i>	93.1	94.1	100.0	92.1	62.4	18.8	9.9	12.9
<i>Journal of Psychopathology and Clinical Science</i>	97.7	97.7	97.7	84.1	50	43.2	45.5	29.5
<i>Journal of Personality and Social Psychology</i>	71.6	98.9	98.9	86.4	37.5	90.9	83.0	73.9
Total	90.4	97.6	99.4	86.6	58.3	50.1	44.7	29.1

reporting of national origin, demographics, and methodological standards. Our results also indicate that the adoption of these methodological practices has been inconsistent. Some recommendations do appear to have been adopted with gusto, for example, over 90% of authors now describe the demographics of their sample in their abstract and justify their sample in text. However, most studies are not pre-registered (with the notable exception of those published in *Journal of Personality and Social Psychology*), and only about half of published research has publicly available data and/or materials. Perhaps most disappointingly, in response to the APA recommendation to provide “constraints on generality” statements for all APA journals (APA, 2018), many authors appear to have written pro forma responses with only half engaging with these limitations meaningfully.

Implications and Future Directions

Many researchers have offered thoughtful strategies for achieving more representative samples. Achieving that growth will require concerted action by authors, editors, and funders. Two broad strategies are often proposed, each with trade-offs. First, representativeness could be achieved by requiring WEIRD researchers to obtain more diverse samples. Muthukrishna et al. (2020), for instance, advocated for stratified sampling and targeted recruitment to address WEIRD biases. Apicella et al. (2020) and Rad et al. (2018) have called for policy changes within funding agencies to incentivize more diverse sampling. Several funding agencies offer models for such calibrated mandates: the National Institutes of Health’s All of Us program shows how policy and infrastructure can advance diversity targets, and federal oversight highlights both feasibility and lingering gaps (Bianchi et al., 2024). While we echo these recommendations in principle, we acknowledge the challenges that lie ahead. In particular, reduced funding for diversity, equity and inclusion initiatives and growing political resistance may hinder progress—especially for U.S.-based researchers. Similar mandates by journals may improve external validity, but without the attached funding, they could raise costs, encourage parachute science, and potentially slow scientific discovery. Undergraduate and online convenience samples remain widespread because they are fast and affordable, especially at institutions with limited funding. Online platforms can, in principle, expand reach beyond traditional student pools and enable stratified recruitment that better mirrors target populations. However, concerns about data quality remain (although may be no worse than student samples; Douglas et al., 2023). Additionally, a move to extra-active research practices (sometimes called parachute science; Stefanoudis et al., 2021) risks WEIRD scholars collecting data from non-WEIRD populations without deep contextual engagement or equitable collaboration.

Alternatively, greater representation could be achieved by encouraging non-WEIRD scholars to publish in APA journals (Rad et al., 2018). The current state of psychological science may present an opportunity to elevate research conducted by non-U.S. scholars working within their own cultural contexts. Calls for APA journals to pursue non-U.S.-based editorial board members and highlight non-American research via special issues (Arnett, 2008; Thalmayer et al., 2021; Webster et al., 2021) echo this idea. On a practical level, free or low-cost translation services (Khelifa et al., 2021) and modern AI translation tools could reduce language barriers and broaden access to non-English scholarship (Steigerwald et al., 2022). However, critics note the risk of reinforcing U.S. journals as arbiters of good science (Adams et al., 2019; Ledgerwood et al., 2024).

Overall, therefore, we see no risk-free solution to the WEIRD problem. But our findings point to a simpler and banal first step: not to invent new standards but to ensure that existing ones are followed. JARS-REC already provides a robust scaffold for transparent, context-sensitive reporting. What is missing is consistent uptake by authors and consistent enforcement by editors and reviewers. We recommend that journals should treat JARS-REC less as “nice-to-have” guidance and more as publishing requirements, requesting explicit, point-by-point compliance at submission and revision and using these criteria in editorial decisions. At the same time, JARS-REC would benefit from clearer distinctions between universal expectations and genuinely conditional items. Many recommendations are framed as “if X, then Y,” which can leave even well-intentioned authors uncertain about applicability. A crisper split, “always report” versus “report when applicable,” with concrete triggers, would help authors plan methods and write-ups and would empower reviewers to enforce compliance without overreach. Finally, we suggest that future assessments continue to treat national origin, demographic representativeness, and reporting standards in combination, not isolation. Tracking any one dimension without the others risks a partial view. Progress in one area can amplify gains elsewhere, whereas weaknesses in any single domain can dilute otherwise meaningful advances.

These recommendations may feel overly modest and pale in comparison to the problem at hand. But in a world where much of the best psychological research still omits basic demographic details, we need to be realistic about where the field is at and meet it where it is. We believe these recommendations raise the floor on routine practice. Once future reviews show that this standard has been reached, more expansive changes can follow.

Limitations and Constraints on Generality

Based on our extended literature review, we feel confident that our estimation of sample “WEIRDness” in APA journals is representative of many U.S.-based psychological journals, across both high- and low-impact outlets, and is likely broadly

indicative of Western psychological science more generally. However, we make no claims beyond this. Psychologists working in non-WEIRD countries do not generally follow APA-style research conventions, and many scholars who focus on local issues or publish in languages other than English are unlikely to submit to APA journals (Adair & Huynh, 2012; Sinha, 1989; Van de Vijver, 2013). Therefore, sample origins, sample demographics, and adherence to reporting standards may look very different in non-U.S. journals. International samples ought to reflect the demographic composition of their own populations, which may lead to a very different diversity target than U.S. research. We also acknowledge that JARS-REC is an American reporting standard. While we view transparent reporting as a universal scientific value, we do not assume that JARS-REC should be uniformly applied or enforced across all international journals or scholarly traditions.

Our research is limited by our coarse operationalization of several JARS-REC reporting practices. JARS-REC is written primarily for authors, who possess intimate knowledge of their research. Because our review spans multiple sub-disciplines, we approached these coding decisions as non-experts. To avoid imposing idiosyncratic or discipline-specific judgments that would (and should) require deep domain expertise (e.g., whether a measurement tool was culturally appropriate), we restricted coding to clear, broadly applicable departures from reporting standards. This approach necessarily overlooks more subtle, graded violations of best practice and as such should be interpreted as identifying only the most obvious instances of noncompliance. Future work would benefit from more granular, severity-weighted coding schemes developed in collaboration with domain experts.

A final limitation concerns the timing of the release and uptake of JARS-REC relative to the publication years examined. JARS-REC was not formally issued until 2023, and therefore many articles published in 2024 likely reflect data that were collected, analyzed, or even drafted prior to the standards’ release, limiting the extent to which full compliance could reasonably be expected. As a result, observed reporting practices in the most recent publication year should be interpreted as an early snapshot of adoption rather than a definitive test of impact.

Conclusion

This review offers a new benchmark for evaluating the state of psychological science. By examining national sample origins, demographic representativeness, and compliance with reporting standards together, we provide a holistic view of the field in 2024. The picture is sobering. Psychology remains disproportionately reliant on U.S. and other WEIRD samples; demographic representativeness is often hard to evaluate because key features go unreported; and adoption of JARS-REC is uneven. In this sense, the benchmark we

establish is low, but therefore of course, the opportunity for growth is substantial.

References

- Adair, J. G., & Huynh, C.-L. (2012). Internationalization of psychological research: Publications and collaborations of the United States and other leading countries. *International Perspectives in Psychology*, 1(4), 252–267. <https://doi.org/10.1037/a0030395>
- Adams, G., Estrada-Villalta, S., Sullivan, D., & Markus, H. R. (2019). The psychology of neoliberalism and the neoliberalism of psychology. *Journal of Social Issues*, 75(1), 189–216. <https://doi.org/10.1111/josi.12305>
- American Psychological Association. (2018). *Journal article reporting standards*. <https://apastyle.apa.org/jars>
- American Psychological Association. (2021). *Apology to people of color for APA's role in promoting, perpetuating, and failing to challenge racism, racial discrimination, and human hierarchy in the United States*. <https://www.apa.org/about/policy/racism-apology>
- American Psychological Association. (2023). *Journal article reporting standards for race, ethnicity, and culture (JARS-REC): Table 1*. <https://apastyle.apa.org/jars/rec-table-1.pdf>
- Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, 41(5), 319–329. <https://doi.org/10.1016/j.evolhumbehav.2020.07.015>
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, 63(7), 602–614. <https://doi.org/10.1037/0003-066X.63.7.602>
- Bauer, L. (2019). Expanding the reach of psychological science. *Journal of Cross-Cultural Psychology*, 50(3), 344–358. <https://doi.org/10.1177/09567976221133816>
- Bauer, L. (2022). Responding to the association for psychological science strategic plan, 2022–2027. *Psychological Science*, 33(1), 3–7. <https://doi.org/10.1177/09567976221133816>
- Berniūnas, R., Beinorius, A., Dranseika, V., Silius, V., & Rimkevičius, P. (2021). The weirdness of belief in free will. *Consciousness and Cognition*, 87, Article 103054. <https://doi.org/10.1016/j.concog.2020.103054>
- Bianchi, D. W., Brennan, P. F., Chiang, M. F., Criswell, L. A., D'Souza, R. N., Gibbons, G. H., Gilman, J. K., Gordon, J. A., Green, E. D., Gregurick, S., Hodes, R. J., Kilmarx, P. H., Koob, G. F., Koroshetz, W. J., Langevin, H. M., Lorsch, J. R., Marrazzo, J. M., Pérez-Stable, E. J., Rathmell, W. K., ... Denny, J. C. (2024). The all of us research program is an opportunity to enhance the diversity of US biomedical research. *Nature Medicine*, 30(2), 330–333. <https://doi.org/10.1038/s41591-023-02744-3>
- Bradford, D. E., DeFalco, A., Perkins, E. R., Carbajal, I., Kwasa, J., Goodman, F. R., Jackson, F., Richardson, L. N. S., Woodley, N., Neuberger, L., Sandoval, J. A., Huang, H. J., & Joyner, K. J. (2024). Whose signals are being amplified? Toward a more equitable clinical psychophysiology. *Clinical Psychological Science*, 12(2), 237–252. <https://doi.org/10.1177/2167702622112117>
- De Man, J., Campbell, L., Tabana, H., & Wouters, E. (2021). The pandemic of online research in times of COVID-19. *BMJ Open*, 11(2), Article e043866. <https://doi.org/10.1136/bmjopen-2020-043866>
- DeJesus, M. A., Callanan, G., Solis, G., & Gelman, S. A. (2019). Generic language in scientific communication. *Proceedings of the National Academy of Sciences*, 116(37) 18370–18377. <https://doi.org/10.1073/pnas.1817706116>
- Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE*, 18(3), Article e0279720. <https://doi.org/10.1371/journal.pone.0279720>
- Forscher, P. S., Basnight-Brown, D. M., Dutra, N., Adetula, A., Silan, M., & IJzerman, H. (2022). Psychological science needs the entire globe, part 3. *Observer*. <https://www.psychologicalscience.org/observer/entire-globe-part-3>
- Gannon, L., Luchetta, T., Rhodes, K., Pardie, L., & Segrist, D. (1992). Sex bias in psychological research: Progress or complacency? *American Psychologist*, 47, 389–396. <https://doi.org/10.1037/0003-066X.47.3.389>
- Graham, S. (1992). “Most of the subjects were White and middle class”: Trends in published research on African Americans in selected APA journals, 1970–1989. *American Psychologist*, 47, 629–639. <https://doi.org/10.1037/0003-066X.47.5.629>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–135. <https://doi.org/10.1017/S0140525X0999152X>
- Jacobs, J. A., & Mizrahi, N. (2020). International representation in US social-science journals. *American Sociologist*, 51(2), 215–239. <https://doi.org/10.1007/s12108-020-09440-6>
- Kahalon, R., Klein, V., Ksenofontov, I., Ullrich, J., & Wright, S. C. (2021). Mentioning the sample's country in the article's title leads to bias in research evaluation. *Social Psychological and Personality Science*, 13(2), 352–361. <https://doi.org/10.1177/19485506211024036>
- Khelifa, R., Amano, T., & Nuñez, M. A. (2021). A solution for breaking the language barrier. *Trends in Ecology & Evolution*, 37(2), 125–128. <https://doi.org/10.1016/j.tree.2021.11.003>
- Kissel, H. A., Lee, G. H., McFarland, S., Berger, D., Enneking, E., Dunham, J., & Brumback, T. (2024). Participant diversity in ACER: 2010–2022. *Alcohol, Clinical & Experimental Research*, 48(6), 1189–1204. <https://doi.org/10.1111/acer.15324>
- Krys, K., de Almeida, I., Wasielec, A., & Vignoles, V. (2024). WEIRD-confucian comparisons: Ongoing cultural biases in psychology's evidence base and some recommendations for improving global representation. *American Psychologist*, 80(2), 247–263. <https://doi.org/10.1037/amp0001298>
- Ledgerwood, A., Lawson, K. M., Kraus, M. W., Vollhardt, J. R., Remedios, J. D., Westberg, D. W., Uskul, A. K., Adetula, A., Leach, C. W., Martinez, J. E., Naumann, L. P., Reddy, G., Chucky Tate, C., Todd, A. R., Weltzien, K., Buchanan, N., González, R., Montilla Doble, L. J., Romero-Canyas, R., ... Zou, L. X. (2024). Disrupting racism and global exclusion in academic publishing: Recommendations and resources for authors, reviewers, and editors. *Collabra: Psychology*, 10(1), Article 121394. <https://doi.org/10.1525/collabra.121394>
- Li, X., Song, D., Han, M., Zhang, Y., & Kizilcec, R. F. (2021). *On the limits of algorithmic prediction across the globe*. arXiv. <https://doi.org/10.48550/arXiv.2103.15212>
- McGorray, E. L., Emery, L. F., Garr-Schultz, A., & Finkel, E. J. (2023). “Mostly White, heterosexual couples”: Examining demographic diversity and reporting practices in relationship science research samples. *Journal of Personality and Social Psychology*, 125(2), 316–344. <https://doi.org/10.1037/pspi0000417>
- Medina, L. D., Torres, S., Gioia, A., Ochoa Lopez, A., Wang, J., & Cirino, P. T. (2020). Reporting of demographic variables in neuropsychological research: An update of O'Bryant et al.'s trends in the current literature. *Journal of the International Neuropsychological Society*, 27(5), 497–507. <https://doi.org/10.1017/S1355617720001083>
- Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, 31(6), 678–701. <https://doi.org/10.1177/0956797620916782>
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, 162, 31–38. <https://doi.org/10.1016/j.jep.2017.04.017>

- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), Article aac4716. <https://doi.org/10.1126/science.aac4716>
- Pedersen, S. L., Lindstrom, R., Powe, P. M., Louie, K., & Escobar-Viera, C. (2022). Lack of representation in psychiatric research: A data-driven example from scientific articles published in 2019 and 2020 in the American journal of psychiatry. *American Journal of Psychiatry*, 179(5), 388–392. <https://doi.org/10.1176/appi.ajp.21070758>
- Pollet, T. V., & Saxton, T. K. (2019). How diverse are the samples used in the journals 'evolution & human behavior' and 'evolutionary psychology'? *Evolutionary Psychological Science*, 5(3), 357–368. <https://doi.org/10.1007/s40806-019-00192-2>
- Protzko, J., Silan, M. A., & Grinschgl, S. (2026). *How to randomly sample the globe: A first step* [Preprint]. PsyArXiv. https://osf.io/preprints/psyarxiv/chqv_v1
- 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*, 115(45), 11401–11405. <https://doi.org/10.1073/pnas.1721165115>
- Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, 15(6), 1295–1309. <https://doi.org/10.1177/1745691620927709>
- Ruiz, E., Beatty, A., & Weyandt, L. (2023). An investigation of diversity in childhood and adolescent antidepressant studies: A systematic review. *Journal of Child and Adolescent Psychopharmacology*, 33(5), 164–175. <https://doi.org/10.1089/cap.2022.0090>
- Singh, L., Barokova, M., Bazhydai, M., Baumgartner, H. A., Franchin, L., Kosie, J. E., Lew-Williams, C., Okyere Omane, P., Reinelt, T., Schuwerk, T., Sheskin, M., Soderstrom, M., Wu, Y., & Frank, M. C. (2024). Tools of the trade: A guide to sociodemographic reporting for researchers, reviewers, and editors. *Journal of Cognition and Development*, 26(3), 354–373. <https://doi.org/10.1080/15248372.2024.2431106>
- Singh, L., & Rajendra, S. J. (2024). Greater attention to socioeconomic status in developmental research can improve the external validity, generalizability, and replicability of developmental science. *Developmental Science*, 27(5), Article e13521. <https://doi.org/10.1111/desc.13521>
- Sinha, D. (1989). Research in psychology in the developing world: An overview. *Psychology and Developing Societies*, 1(1), 105–126. <https://doi.org/10.1177/097133368900100109>
- Solomonov, N., Chen, S. Z., Briskin, E. A., Castonguay, L., Krause, M., McMain, S., Duggal, C., Youn, S. J., Lorenzo-Luaces, L., & Barber, J. P. (2025). A first look at diversity gaps in psychotherapy research publications and representation. *Psychotherapy Research*, 35, 1275–1286. <https://doi.org/10.1080/10503307.2024.2428693>
- Stefanoudis, P. V., Licuanan, W. Y., Morrison, T. H., Talma, S., Veitayaki, J., & Woodall, L. C. (2021). Turning the tide of parachute science. *Current Biology*, 31(4), R184–R185. <https://doi.org/10.1016/j.cub.2021.01.029>
- Steigerwald, E. C., Ramírez-Castañeda, V., Brandt, D. Y. C., Báldi, A., Shapiro, J. T., Bowker, L., & Tarvin, R. D. (2022). Overcoming language barriers in academia: Machine translation tools and a vision for a multilingual future. *Bioscience*, 72(10), 988–998. <https://doi.org/10.1093/biosci/biac062>
- Thalmayer, A. G., Toscanelli, C., & Arnett, J. J. (2021). The neglected 95% revisited: Is American psychology becoming less American? *American Psychologist*, 76(1), 116–129. <https://doi.org/10.1037/amp0000622>
- Van de Vijver, F. J. (2013). Contributions of internationalization to psychology: Toward a global and inclusive discipline. *American Psychologist*, 68(8), 761–770. <https://doi.org/10.1037/a0033762>
- Veritas Health Innovation. (n.d.). *Covidence* [Systematic review software]. <https://www.covidence.org/>
- Webster, G. D., Mahar, E. A., & Wongsomboon, V. (2021). American psychology is becoming more international, but too slowly: Comment on Thalmayer et al. (2020). *American Psychologist*, 76(5), 802–805. <https://doi.org/10.1037/amp0000747>

Received May 13, 2025

Revision received January 11, 2026

Accepted January 14, 2026 ■