Framework for Open and Reproducible Research Training | FORRT

Tracking and mainstreaming replications in the social, cognitive, and behavioral sciences

Helena Hartmann, Department for Neurology and Center for Translational Neuro- and Behavioral Sciences (C-TNBS), University Hospital Essen, Germany
helena.hartmann@uk-essen.de | 0000-0002-1331-6683

Flavio Azevedo, Department of Interdisciplinary Social Science, Utrecht University
f.a.azevedo@uu.nl | 0000-0001-9000-8513

Lukas Röseler, Münster Center for Open Science, University of Münster, Germany
Email | Hyperlinked ORCID

Lukas Wallrich, Birkbeck Business School, University of London, United Kingdom
l.wallrich@bbk.ac.uk | 0000-0003-2121-5177

Alaa Aldoh, Department of Psychology, University of Amsterdam, the Netherlands
a.aldoh@uva.nl | 0000-0003-1988-0661

Mahmoud M. Elsherif, Department of Psychology and Vision Sciences, University of Leicester, United Kingdom; Department of Psychology, University of Birmingham, United Kingdom
mahmoud.medhat.elsherif@gmail.com | 0000-0002-0540-3998

Meng Liu, School of English and International Studies, Beijing Foreign Studies University, China
Email | Hyperlinked ORCID

Rían O Mahoney, Arb Research, Czechia
Email | 0009-0007-4939-0766

Zoran Pavlović, Department of Psychology, Faculty of Philosophy, University of Belgrade, Serbia
zoran.pavlovic@f.bg.ac.rs | 0000-0002-9231-5100

Aleksandrina Skvortsova, Faculty of Social Sciences, Leiden University, the Netherlands
a.skvortsova@fsw.leidenuniv.nl | 0000-0003-0512-0792

Valeria Agostini, School of Psychology, University of Birmingham, United Kingdom; Centre for Developmental Science, University of Birmingham, United Kingdom
valeria.agostini4@gmail.com | 0000-0003-0314-2998

Laurence Aitchison, University of Bristol, United Kingdom
Email | Hyperlinked ORCID

Ali H. Al-Hoorie, Jubail English Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Saudi Arabia

Email | Hyperlinked ORCID

Samuel Alarie, Department of Psychology, University of Montreal, Canada

Email | Hyperlinked ORCID

Nihan Albayrak, Department of Psychology, Boğaziçi University, Türkiye; Department of Psychological and Behavioural Science, London School of Economics and Political Science, United Kingdom

nihanalb@bogazici.edu.tr | 0000-0003-3412-4311

Nora Ammann, ARIA, United Kingdom

Email | Hyperlinked ORCID

Farid Anvari, Institute of Psychology, University of Bern, Switzerland

faridanvari.fa@gmail.com | 0000-0002-5806-5654

Patricia Arriaga, Department of Psychology, Iscte-University Institute of Lisbon, Cis-Iscte, Portugal

patricia.arriaga@iscte-iul.pt | 0000-0001-5766-0489

Bradley J. Baker, Department of Sport, Tourism and Hospitality Management, Temple University, United States

bradley.baker@temple.edu | 0000-0002-1697-4198

Balazs Aczel, Institute of Psychology, ELTE, Eotvos Lorand University, Hungary

Email | Hyperlinked ORCID

David J. Bauer, Viterbo University, United States

Email | Hyperlinked ORCID

Julia Beitner, Department of Psychology, Goethe University Frankfurt, Germany; Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany

Email | Hyperlinked ORCID

Anabel Belaus, Instituto de Investigaciones Psicológicas (IIPsi) - CONICET - UNC, Costa Rica

Email | Hyperlinked ORCID

Hetvi Bhatt, Enable, United Kingdom

Email | Hyperlinked ORCID

Cameron Brick, Department of Psychology, University of Amsterdam, the Netherlands

Email | Hyperlinked ORCID

Hilmar Brohmer, Department of Psychology, University of Graz, Austria; University of Bern, Switzerland

hilmar.brohmer@uni-graz.at | 0000-0001-7763-4229

Benjamin Brummernhenrich, Institute of Psychology for Education, University of Münster, Germany

Email | Hyperlinked ORCID

Emily Budd, Behavioral Sciences, Webster University, Netherlands
Email | 0009-0004-9021-1035

Anya Butler, Department of Experimental Psychology, University of Oxford, United Kingdom
Email | Hyperlinked ORCID

Subramanya P. Chandrashekar, Department of Psychology, Norwegian University of Science and Technology, Norway
Email | Hyperlinked ORCID

Sau-Chin Chen, Department of Human Development and Psychology, Tzu-Chi University, Taiwan
pmsp96@gmail.com | 0000-0001-6092-6049

Kai Li Chung, University of Reading, United Kingdom
Email | Hyperlinked ORCID

Jamie P. Cockcroft, Department of Psychology, University of York, United Kingdom
jamie.cockcroft@york.ac.uk | 0000-0002-0637-8851

Paul Crowe, Arb Research, United Kingdom
paulcrowe1@protonmail.com | 0009-0004-1030-5198

Oliver Deane, School of Computer Science, University of Bristol, United Kingdom
Email | Hyperlinked ORCID

Veronica Diveica, Montreal Neurological Institute, McGill University, Canada
veronicadiveica@gmail.com | 0000-0002-5696-8200

Andis Draguns, UL Institute of Mathematics and Computer Science, Latvia
Email | Hyperlinked ORCID

Sam Enright, School of Philosophy, Psychology, & Language Sciences, University of Edinburgh, United Kingdom
Email | Hyperlinked ORCID

Thomas R. Evans, School of Human Sciences and Institute for Lifecourse Development, University of Greenwich, United Kingdom
Email | Hyperlinked ORCID

Anna Exner, Faculty of Psychology, Ruhr University Bochum, Germany
Email | Hyperlinked ORCID

Gilad Feldman, Department of Banking and Finance, University of Innsbruck, Innsbruck, Austria
giladfel@gmail.com | 0000-0003-2812-6599

Adrien Fillon, LAPSCO, CNRS UMR 6024, Université Clermont-Auvergne, France
adrien.fillon@uca.fr | 0000-0001-8324-2715

Felipe F. Vieira, Department of Data Analysis, Faculty of Psychology and Educational Sciences, Ghent University, Belgium

Email | Hyperlinked ORCID

Joris Frese, Department of Political and Social Sciences, European University Institute, Italy
joris.frese@eui.eu | 0000-0002-5871-997X

Max C.D. Gattie, University of Manchester, United Kingdom
Email | Hyperlinked ORCID

Vaitsa Giannouli, School of Medicine, Aristotle University of Thessaloniki, Greece; Department of Psychology, University of Western Macedonia, Greece
giannouliv@hotmail.com | 0000-0003-2176-8986

Biljana Gjoneska, Macedonian Academy of Sciences and Arts, North Macedonia
Email | Hyperlinked ORCID

Amélie Gourdon-Kanhukamwe, King's College London, United Kingdom
Email | Hyperlinked ORCID

Christopher J. Graham, Department of Education, Training and Assessment, Royal College of Physicians of Edinburgh, United Kingdom
christopherjgraham93@gmail.com | 0000-0002-1144-7970

Jason Hausenloy, Department of Mathematics, University of California at Berkeley, United States
Email | Hyperlinked ORCID

Alina Herderich, IDea Lab, University of Graz, Austria
Email | Hyperlinked ORCID

Zlatomira G. Ilchovska, Department of Psychology, University of York, United Kingdom; School of Psychology, University of Birmingham, United Kingdom; School of Psychology, University of Nottingham, United Kingdom
z.g.ilchovska@bham.ac.uk | 0000-0001-6682-9952

Hirotaka Imada, Department of Psychology, Royal Holloway, University of London, United Kingdom
Hirotaka.Imada@rhul.ac.uk | 0000-0003-3604-4155

Kamil Izydorczak, Wrocław Faculty of Psychology, SWPS University of Social Sciences and Humanities, Poland
Email | Hyperlinked ORCID

Alma Jeftić, Swiss Centre for Affective Sciences, University of Geneva, Switzerland; Peace Research Institute, International Christian University, Japan
alma.jeftic@gmail.com | 0000-0002-9285-2061

Tamara Kalandadze, Department of Education, ICT and Learning, Østfold University College, Norway
Email | Hyperlinked ORCID

Alexandros Kastrinogiannis, Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Email | Hyperlinked ORCID

Maren Klingelhöfer-Jens, Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany

Email | Hyperlinked ORCID

Halil E. Kocalar, Department of Psychological Counseling and Guidance, Muğla Sıtkı Koçman University, Turkey

hemrekocalar@mu.edu.tr | 0000-0002-7299-162X

Lina Koppel, Division of Economics, Department of Management and Engineering, Linköping University, Sweden

Email | Hyperlinked ORCID

Alina Koppold, University of Bielefeld, Biological Psychology and Cognitive Neuroscience, Bielefeld, Germany; Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany

alina.koppold@uni-bielefeld.de | 0000-0002-3164-3389

Max Korbmacher, NeuroSys-Med group, Department of Neurology, Haukeland University Hospital, Norway; Department of Radiography, Faculty of Health and Functioning, Western Norway University of Applied Sciences, Norway

Email | Hyperlinked ORCID

Annalise LaPlume, Trent University, Department of Psychology, Canada

annaliselaplume@trentu.ca | 0000 0001 6725 3270

Jaeho Lee, Arb Research, United States

Email | Hyperlinked ORCID

Nigel Mantou Lou, Department of Psychology, University of Victoria, Canada

nigellou@uvic.ca | 0000-0003-1363-833X

Dermot Lynott, Department of Psychology, Maynooth University, Ireland

dermot.lynott@mu.ie | 0000-0001-7338-0567

Sharan Maiya, Language Technology Lab, University of Cambridge, United Kingdom

sm2783@cam.ac.uk | 0009-0003-3658-9873

David McSharry, Kreoh, Ireland

Email | Hyperlinked ORCID

Kimberly L. Meidenbauer, Department of Psychology, Washington State University, United States

k.meidenbauer@wsu.edu | 0000-0001-9135-6130

Maria Meier, University of Konstanz, Germany; University Psychiatric Clinics Basel, University of Basel, Switzerland

maria.meier@uni-konstanz.de | 0000-0002-1655-5479

Leticia Micheli, Institute of Psychology, Leiden University, the Netherlands
l.rettore.micheli@fsw.leidenuniv.nl | 0000-0003-0066-8222

Maria Montefinese, Department of Developmental Psychology and Socialisation, University of Padova, Italy
maria.montefinese@unipd.it | 0000-0002-7685-1034

David Moreau, School of Psychology and Centre for Brain Research, University of Auckland, New Zealand
d.moreau@auckland.ac.nz | 0000-0002-1957-1941

Nadja K. Moser, Department of Neurosurgery, Charité University Medicine Berlin, Germany
Email | Hyperlinked ORCID

Kellen Mrkva, Hankamer School of Business, Baylor University, United States
Email | Hyperlinked ORCID

Niyatee Narkar, Department of Psychology, University of Guelph, Canada
nnarkar@uoguelph.ca | 0009-0006-5219-8305

Monika Nemcova, Graduate School of Systemic Neurosciences, Ludwig Maximilian University, Germany
Email | Hyperlinked ORCID

Yvonne Oberholzer, University of Zurich, Jacobs Center for Productive Youth Development
Email | Hyperlinked ORCID

Aoife O'Mahony, School of Public Health, University College Cork, Ireland
Email | Hyperlinked ORCID

Julian Packheiser, Social Neuroscience, Faculty of Medicine, Ruhr University Bochum, Germany; Research Center One Health Ruhr of the University Alliance Ruhr, Ruhr University Bochum, Germany
Email | Hyperlinked ORCID

Shubham Pandey, Department of Humanities and Social Science, Indian Institute of Technology Bombay, India; Institute of Psychology, Osnabrueck University, Germany
shubham.cogsci@gmail.com | 0000-0002-5591-8551

Charlotte R. Pennington, School of Psychology, Aston University, Birmingham, United Kingdom
Email | Hyperlinked ORCID

Merle-Marie Pittelkow, QUEST Center for Responsible Research, Germany; Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Germany
Email | Hyperlinked ORCID

Willemijn Plomp, Leiden University, the Netherlands
Email | Hyperlinked ORCID

Paul E. Plonski, Swarthmore College, United States
Email | Hyperlinked ORCID

Ekaterina Pronizius, Department of Cognition, Emotion, and Methods in Psychology, University of Vienna, Austria

ekaterina.pronizius@univie.ac.at | 0000-0003-1446-196X

Andrew Adrian Yu Pua, Carlos L. Tiu School of Economics, De La Salle University, Philippines; School of Business, Economics and Information Systems, University of Passau, Germany

andrewypua@outlook.com | 0000-0002-2225-5245

Robert Reason, Arb Research, United Kingdom; School of Psychological Science, University of Bristol, United Kingdom

Email | Hyperlinked ORCID

Elena Richert, Department of Psychology, Reykjavik University, Iceland; Reykjavik University Sleep Institute, Reykjavik University, Iceland; Department of Technical Physics, University of Eastern Finland, Finland

elenar@ru.is | 0000-0003-0919-4879

Jan P. Röer, Department of Psychology and Psychotherapy, Witten/Herdecke University, Germany

Email | Hyperlinked ORCID

Robert M. Ross, School of Psychological Sciences and School of Humanities, Macquarie University, Australia

Email | Hyperlinked ORCID

Kathleen Schmidt, Department of Psychology, Ashland University, United States

Email | Hyperlinked ORCID

Nuño Sempere, Sentinel Research, Paraguay

Email | Hyperlinked ORCID

Matthias F. J. Sperl, Department of Clinical Psychology and Psychotherapy, University of Siegen, Germany; Department of Clinical Psychology and Psychotherapy, University of Giessen, Germany; Center for Mind, Brain and Behavior, Universities of Marburg and Giessen (Research Campus Central Hessen), Germany

Email | Hyperlinked ORCID

Jeffrey R. Stevens, Department of Psychology, University of Nebraska-Lincoln, United States

Email | Hyperlinked ORCID

Alvin W. M. Tan, Department of Psychology, Stanford University, United States

tanawm@stanford.edu | 0000-0001-5551-7507

Marina Tiulpakova, RWTH Aachen University, Germany

marina.tiulpakova@rwth-aachen.de | 0009-0006-1346-6741

Aleksandra Tołopiło, Center for Research on Biological Basis of Social Behavior, SWPS University, Poland

atolopilo@swps.edu.pl | 0000-0002-2518-6759

Burak Tunca, School of Economics and Management, Lund University, Sweden

Email | Hyperlinked ORCID

Wolf Vanpaemel, Faculty of Psychology and Educational Sciences, KU Leuven, Belgium
wolf.vanpaemel@kuleuven.be | 0000-0002-5855-3885

Leigh Ann Vaughn, Department of Psychology, Ithaca College, United States
Email | Hyperlinked ORCID

Steven Verheyen, Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, the Netherlands
verheyen@essb.eur.nl | 0000-0002-6778-6744

Julia Wolska, School of Psychology, Manchester Metropolitan University, United Kingdom
j.wolska@mmu.ac.uk | 0000-0001-8675-4388

Siu Kit Yeung, Department of Psychology, Chinese University of Hong Kong, China
Email | Hyperlinked ORCID

Mirela Zaneva, Christ Church College, University of Oxford, United Kingdom
Email | Hyperlinked ORCID

Gavin Leech, Leverhulme Centre for the Future of Intelligence, University of Cambridge, United Kingdom
Email | Hyperlinked ORCID

Please cite this work as:

Hartmann, H., Azevedo, F., Röseler, L., Wallrich, L., Aldoh, A., Elsherif, M. M. et al. (2025, November). Tracking and mainstreaming replications in the social, cognitive, and behavioral sciences. *Preprint.*

Correspondence:
Correspondence concerning this article should be directed to Flavio Azevedo: f.a.azevedo@uu.nl.

Author Contributions:
Authorship order was set as follows: Project coordinators (HH, FA, LR, & LW), project managers in alphabetical order (AA, MME, ML, ROM, ZP, & AS), project contributors in alphabetical order, last author (GL). The first version of the manuscript was drafted by FA, HH, LR, & LW. All authors gave feedback, revised, and approved the final manuscript before submission.

Funding
HH was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project ID 422744262 – TRR 289). FA, LW, and LR were supported by the Center for Open Science. MME was supported by the Leverhulme Trust Early Career

Award (RM56G0344/44143). MCDG was supported by the Economic and Social Research Council. GL was supported by the UK Research and Innovation (UKRI) Centre for Doctoral Training in Interactive Artificial Intelligence Award (EP/S022937/1). VD was supported by a Postdoctoral Fellowship from Fonds de Recherche du Quebec.

Data Accessibility Statement

The data pertaining to this work is accessible at https://forrt.org/apps/fred_explorer.html and is described in Röseler et al. (2024). The dataset is accessible at https://osf.io/9r62x/. The R package can be found at https://forrt.org/FReD/index.html and is described in more detail in Röseler et al. (2025).

Declarations of Conflicting Interests

The authors declare the following competing interests: We all are or were in the past members of the Framework for Open and Reproducible Research Training (FORRT; forrt.org). LR is Editor in Chief of the *Replication Research* journal that is dependent on researchers conducting replications. FA and LW are Senior Editors there. SV and TRE is an Associate Editor there.

Acknowledgments

This resource and work was created as part of the Framework for Open and Reproducible Research Training (FORRT). The FORRT Replication Hub with all its resources can be accessed here. We acknowledge work by the following people: ….

Ethics Approval:

Ethics approval statement is not applicable.

Patient Consent Statement:

Consent Statement is not applicable.

Abstract

Replicability is a cornerstone of scientific progress. Yet, replications are often undervalued, and are sometimes seen as redundant, unimportant, or lacking novelty. This impedes their broader adoption in research and beyond. In response, the credibility revolution calls for slower, more deliberate science and greater responsiveness to fallibility. In this perspective piece, we argue that (a) replications are essential for validating scientific claims, (b) replications need to be made more visible, recognized, and integrated into research and educational practices, and (c) we can change the way we view and judge replication results. We propose a framework where replication studies can be systematically tracked and normalized through the Replication Hub as part of the Framework for Open and Reproducible Research Training (FORRT) initiative, with the goal of enhancing the visibility, integration, and cumulative impact of replication research across disciplines.

Keywords: Metascience; Open scholarship; Open science; Replicability; Replications; Reproducibility; Replication Crisis

Introduction

Replication is the process by which researchers test whether the same claims, in identical, similar, or varying contexts reach conclusions consistent with those of the original study (Parsons et al., 2022). Replications are a cornerstone of empirical research, where independent sources contribute cumulative evidence to support or refute a given claim. However, recent metascientific studies across scientific fields, particularly those in social, cognitive, and behavioral sciences, have found that many prominent findings ‘fail’ to replicate, that is, their results do not converge with those from the target studies and challenge the credibility of previous scientific claims (Brodeur et al., 2024; Ioannidis, 2005; Nosek & Errington, 2020).

Even when studies do replicate, the observed effects are often much smaller (Patil et al., 2016), averaging half of the originally reported effect sizes (Camerer et al., 2018; Open Science Collaboration, 2015). Such patterns appear across fields, including psychology (Klein et al., 2014, 2018; Wagenmakers et al., 2011), medicine (Hope et al., 2021), biology (Begley & Ellis, 2012; Errington et al., 2021a; 2021b), economics (Camerer et al., 2016), and neuroscience (Boekel et al., 2015). This has been coined as a ‘replication crisis’, raising concerns regarding the robustness of scientific knowledge and challenging the validity of decades of research. In response, this so-called crisis has given rise to a grassroots open science reform movement and the emergence of the field of metascience, that is, research on how science is conducted.

Despite this movement and evidence that many studies do not replicate, replication attempts are still rare (Ankel-Peters et al., 2023; Clarke et al., 2024; Hardwicke et al., 2021; Kamermans et al., 2024; Kelly, 2006; Kobrock et al., 2023; Makel et al., 2012; Makel & Plucker, 2014; Marsden et al., 2018; Martin and Clarke, 2017; McNeeley & Warner, 2015; Mueller-Langer et al., 2019). Moreover, journals are often unwilling to publish replication studies, which compromises our ability to build robust bodies of evidence to inform policy and practice. It also highlights that replications are not yet given due recognition, particularly by journal editors, funders, policymakers, and even researchers themselves. In contrast, novel results are often less scrutinized regarding reproducibility, published more readily and cited more frequently (Scheel et al., 2021; Serra-Garcia & Gneezy, 2021).

Contrasting the underappreciation of replications, especially those that challenge long-established original findings, we argue for replications to be seen as a critical resource for designing, conducting, and interpreting research. Viewing the ‘replication crisis’ as a ‘credibility revolution’ (Korbmacher et al., 2023; Vazire, 2018) and an opportunity (Munafò et al., 2022), we are not alone in calling for slower, more deliberate science (Alleva, 2006; Frith, 2020; Owens, 2013; Stengers, 2017) and greater responsiveness to fallibility (Bishop, 2018). In the following sections, we discuss (a) the essential role of replications in science, (b) the need for their increased visibility and recognition through systematic tracking, (c) changing the way we judge replication results, and (d) future directions for replication practices in professional and educational contexts.

The need for replications

There is a common and longstanding narrative of science being built on replications, but recently they have been heralded as a key tool for 'saving science' (Edlund et al., 2021), for example, by ensuring the reliability and validity of scientific findings, strengthening confidence in research outcomes, and identifying potential biases in original studies. Two fundamental aspects of science make replications indispensable: First, given the probabilistic nature of research and the myriad contextual and random factors affecting outcomes, no single study can be conclusive — including in the social, cognitive, and behavioral sciences. Second, science should be self-correcting, cumulative, and incremental, with progress building on prior work.

Despite this, current scientific practice often prioritizes novelty over replication and treats individual findings as definitive rather than part of a larger evolving picture (for example, see Owen, 2013). The credibility revolution has underscored the dangers of prioritizing flashy and unexpected results over robustness. For example, research on social priming appeared so convincing that Nobel laureate Daniel Kahneman dedicated a chapter to it in his bestselling book *Thinking, Fast and Slow* (Kahneman, 2011), and others began exploring its applications in business and health interventions (Papies et al., 2016). However, once preregistered replications were conducted more systematically, multiple independent research teams failed to replicate the originally reported social priming effects (for example, Mac Giolla et al., 2024), and the field became emblematic of the concerns surrounding research integrity (Chivers, 2019; Kahneman, 2012; Leys, 2024; Schimmack et al., 2017; Yong, 2012).

Direct replications are a crucial safeguard against the immense resource use of building a literature on false positive findings (Zwaan et al., 2018). By recreating studies with highly similar or identical methods and sample characteristics, direct replications help to identify which findings are reliable, and, therefore, worth expanding upon (as opposed to previously more common *conceptual* replications, that is, attempts to identify the same effects but often including differences in sample, research designs, measurement approaches and/or analysis pathways; LeBel et al., 2018; Parsons et al., 2022). Given the regular occurrence of false positive results, significantly amplified by publication bias and questionable research practices (John et al., 2012; Nagy et al., 2025; Simmons et al., 2011), multiple and direct replications are essential for science, specifically for ensuring reliable, unbiased results, cumulative knowledge generation, and strengthening scientific rigor. Although replications are not immune to errors, there has been an increasing effort to conduct them with higher statistical power than the original studies and to employ preregistered study designs and analysis plans, thus providing stronger evidence for the robustness of key findings (Hedges & Schauer, 2019). On the other hand, replications per se are not 'better' than original studies and each study needs to be judged on its own merits.

Beyond verifying the existence of an effect, especially when science moves towards application, it is crucial to estimate accurate effect sizes to determine practical significance — whether an effect is meaningful enough to act on or not (Anvari et al., 2023; Peetz et al., 2024). For example, knowing that a medication reliably increases sleep by eight minutes is vital in judging its usefulness, effectiveness, and overall cost-benefit (Ferracioli-Oda et al., 2013). Effect size estimates require large samples for precision, and biases in the literature often exaggerate effect sizes (for example publication bias; Schäfer & Schwarz, 2019). Therefore, multiple or large-scale replications are necessary to provide more precise, reliable, and generalizable estimates of true effects (Forscher et al., 2023; Hunter, 2001; Tiokhin et al., 2019; but see also Ghai et al., 2024).

In addition to simply corroborating or challenging original claims, replications also help identify 'boundary' conditions that affect the presence and/or magnitude of effects (Bauernfeind, 1968). When replication results challenge original studies, authors often cite contextual factors to explain failures. While this may seem like deflecting criticism, it presents opportunities to test such potential factors and generate further hypotheses (Zwaan et al., 2018). If original authors more widely adopt statements of constraints on generality (Simons et al., 2017; Zhu et al., 2025), this process can accelerate. This is particularly crucial when moving beyond limited contexts (for example, WEIRD = Western, Educated, Industrialized, Rich, and Democratic; Ghai et al., 2024; Henrich et al., 2010), where findings need to be tested for broader applicability and generalizability (such as different locations or resources).

While not sufficient on their own, direct or close replications play an integral role for scientific progress, as they ensure that the core effects hold under similar circumstances. Conceptual replications are a crucial next step. In contrast to direct replications, they deliberately vary contextual or methodological features and thus allow assessing the robustness and generalizability of an effect. For example, Tunç and Tunç (2023) propose the Systematic Replications Framework (SRF) to design a pre-planned series of systematically interlinked close and conceptual replications (see also Hüffmeier et al., 2016).

Therefore, studies might ideally first reproduce and replicate previous findings before strategically adding conditions or measures that can provide further insights (that is, *extensions,* cf. constructive replications in Hüffmeier et al., 2016). While many studies replicate main effects before testing interactions, moderators, or mediators, these tests are rarely labeled as replications and often deviate from original study protocols (for example, Röseler et al., 2024; Urminsky et al., 2024). This lack of consistency in naming and methods limits the accumulation of evidence and the tracing of 'failed' replications. One reason for this is that the latter usually remain unpublished. Importantly, 70% of researchers have reported failing to replicate findings at least once (Baker, 2016). Yet, the low publication rate of replications suggests many of these unsuccessful attempts are left in the metaphorical 'file drawer' (Rosenthal, 1979), and are never published, keeping potentially flawed research lines alive (Ferguson, 2012).

Taken together, these developments highlight why replicating results and making these results more visible are fundamental for reliable, trustable science (Anvari & Lakens, 2018; Wingen et al., 2020). While fostering a replication culture is vital, choosing replication targets comes with its own challenges (see Field et al., 2019, for a similar perspective). Relying on findings from a single sample is risky, and can lead to a waste of research resources and a loss of trust in science (Isager et al., 2024). In this context, fostering a culture of replication offers benefits beyond merely assessing individual claims (Feldman, 2025). The expectation of future replication can, for example, improve reporting practices, making research more reproducible, reducing errors, and potentially even preventing fraud (Soderberg et al., 2021). Shifting incentives toward replicable findings rather than novelty (or at least giving them equal attention) could enhance scientific rigor. Additionally, replications often generate open materials and code, further facilitating future research. Replications are also increasingly integrated into research training, offering valuable opportunities for students and fostering international collaborations through large-scale, multi-lab studies (for example, the Many Labs projects; Quintana, 2021; Wagge et al., 2019).

Despite these promises, tracking replication outcomes remains challenging due to the limited availability of comprehensive databases of replication studies, making it difficult to accurately estimate their prevalence. Existing estimates suggest that between 0.2% and 5% of published studies in psychology are replications, with even lower estimated rates in other fields (see Clarke et al., 2023 for results on the 100 highest impact psychology journals from 2010-2021; Hardwicke et al., 2021; Makel et al., 2012 for results on the 100 psychology journals with the highest 5-year impact factors since 1900).

Tracking replications systematically

Practical solutions are essential to shift replication studies from a niche effort to a mainstream scientific practice. To achieve the aim of making replications more mainstream and visible, we created a comprehensive database of replications as a resource for research and teaching. At present, the FORRT Replication Database (FReD) contains a large index of original studies, their replications, and their raw statistics and effect sizes (n = 1,118 original articles from 151 different journals as of 2025-02-11; see Table 1 for an overview; Chawla, 2024; Röseler et al., 2024). With over 160 researchers at the time of publication having contributed to the project since its conception in April 2022, we aim for this resource to be a living, community-driven solution to collecting, updating, and disseminating replications, as well as capturing a broad range of past results and aggregating knowledge to assist both research and teaching (as done routinely for living meta-analyses; Nikolakopoulou et al., 2018; and community-augmented meta-analyses; Burgard et al., 2022).

Table 1. Descriptive data of the FORRT Replication Database as of February 2025.

Type	**Quantity**
Original References	1,118
Replication References	1,137
Number of Journals (Original References)	151

This database is further embedded within the FORRT Replication Hub (https://forrt.org/replication-hub/), a comprehensive and living resource where authors, reviewers, educators, and editors can log and access replication studies. In this hub, the FORRT Replication Database (FReD) hosts a strong infrastructure providing information about replications, and features a large and growing list of resources: 1) The FReD Explorer is a database of original studies and their replications (see Figure 1); 2) The FReD Reference Annotator is a tool to check reference lists for replications listed in that database (see Figure 2); 3) A list of large-scale replication projects. This centralized resource facilitates finding replications in the first place, stimulates discussions amongst scholars from different disciplines, and eases accessibility and integration into scholarly workflows. Moreover, it facilitates the citation of replications alongside original studies, making them easier to incorporate into future research and education. In addition, the database is available to meta-scientists for integration into other platforms, and forms the foundation for dissemination tools such as browser plug-ins that are currently under development.

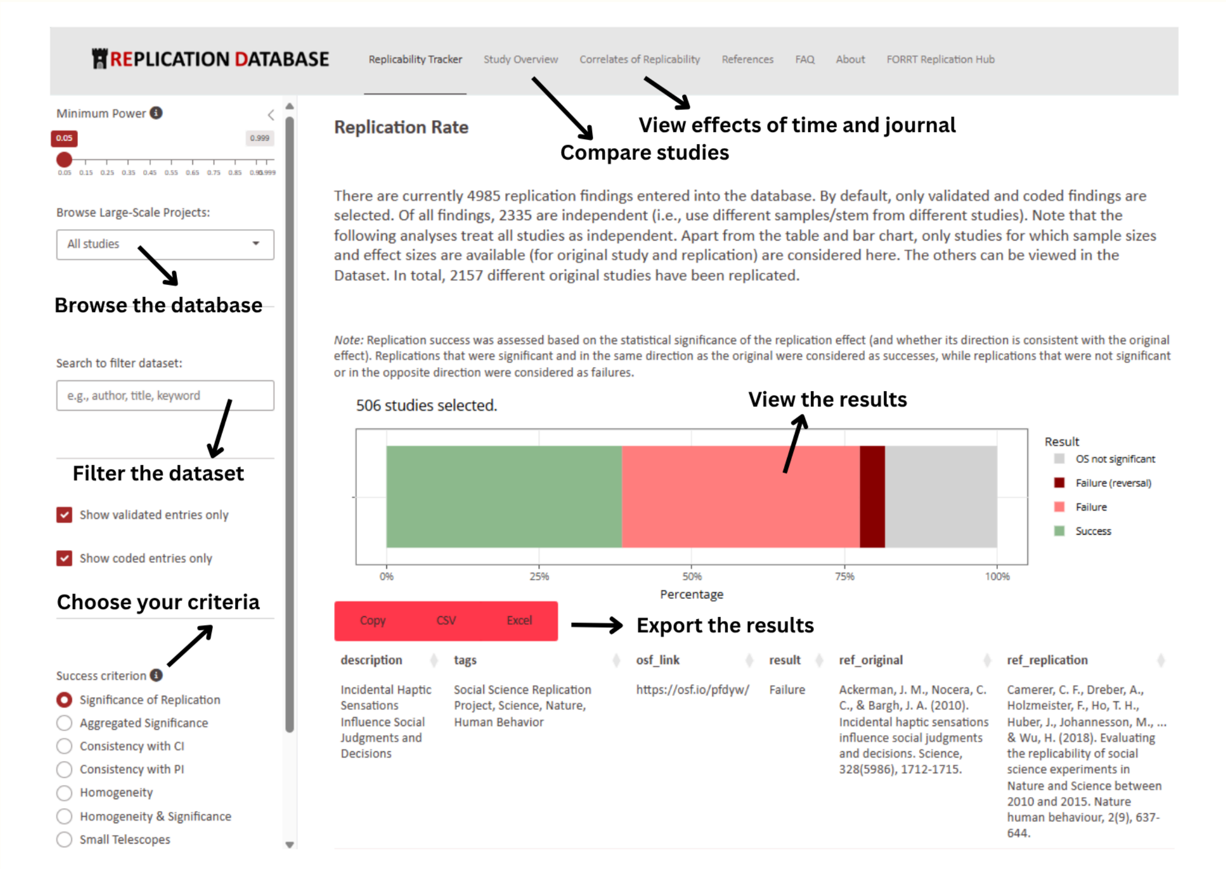

Figure 1. The FORRT Replication Database (FReD) Explorer (as of August 2025), which includes an automated summary of selected replications and success rates as well as filtering options for minimum replication power, project type (for example, Many Labs, Registered Reports, individual replications), validation status, and replication success criterion.

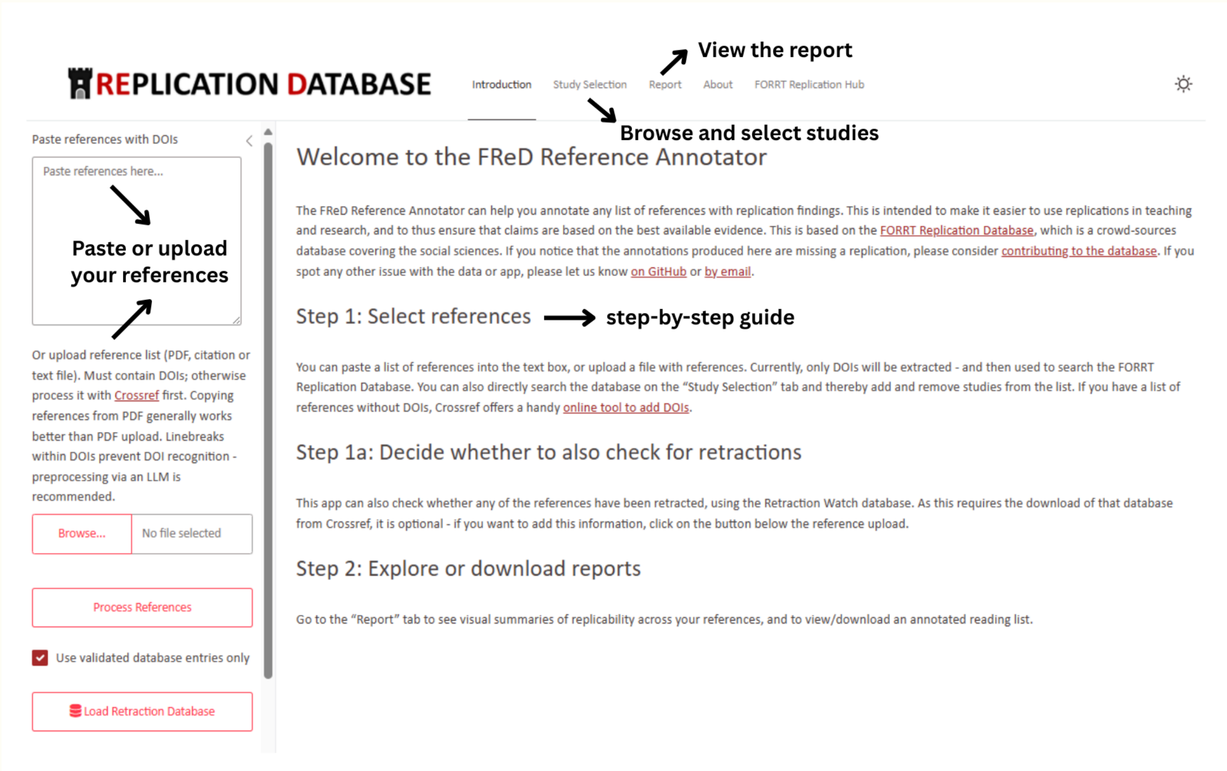

Figure 2. The FORRT Replication Database (FReD) Annotator (as of August 2025), which includes the ability to check reference and reading lists in order to identify pertinent replication studies.

Historically, the initial version of the database involved gathering instances of replication failures and successes from various sources, including scientific mailing lists, blogs, and social media platforms (LeBel et al., 2018; see also the FORRT Replications & Reversals project; https://forrt.org/reversals/), to enable teachers and educators to include replications more readily in their curriculum. Subsequently, participating volunteers at FORRT contributed information about replication studies from their respective subfields over multiple years and at various hackathons starting in 2018. These studies were then sorted by the effect or claim, and collected information was recorded for each study, which included the citation of the study, as well as the study design, sample sizes, and effect sizes of both the original and replication work.

This database does have some limitations worth noting. Due to the self-selected sample of studies, we explicitly refrained in this manuscript from presenting simple summaries or inferential statements about fields or subfields based on the database alone. Neither is the resource an exhaustive list of replications or even 'failed' replications, as our initial collection process was biased towards famous original works, and surprising replication results may have been identified more easily by contributors in their respective fields. However, new evidence about many effects is added weekly (still largely volunteer-driven and -dependent, with recent financial support from the Center for Open Science) and we are making efforts to safeguard against such selection biases, which will become less and less of a problem with the popularity of our database. Lastly, our own effort to collate quantitative features of replications has its own subjectivity and researcher degrees of freedom. For instance, the original papers often consist of a series of experiments testing several hypotheses and often reporting several statistical analyses, and volunteers may choose to add only experiments or effects they subjectively consider to be most pertinent.

Making replications more visible

Once researchers begin to conduct more replications, the next challenge is ensuring that replications become a more easily accessible, valued, and normative part of scientific practice. Key interested parties, including researchers, journals, funders, and policymakers, play critical roles in embedding replication into the research culture (Evans et al., 2022).

The full value of replications can only be realized if they are systematically incorporated into grant applications, publications, and educational curricula. This integration will increase the citations and recognition of replication in the academic community. For example, educators could include replication studies in their syllabi to ensure students build their work on a solid foundation of robust findings and let their students conduct their own small-scale replications (Frank & Saxe, 2012; Hawkins et al., 2018; Kohrs et al., 2023; see Pennington, 2024, for a reflection).

Our own bottom-up efforts need to be reinforced by top-down support from journals and funders. This includes providing explicit incentives for replication research, establishing more replication-specific journals, and revising manuscript evaluation criteria to reduce the emphasis on novelty and innovation during the review process. For instance, a few journals already explicitly invite and publish replication studies, and initiatives like the Registered Reports format incentivize replications by reviewing study designs before data collection, thus reducing publication bias (Scheel et al., 2021). Especially high-impact journals introducing a dedicated article type to promote replications and regularly releasing special issues on this topic would set great examples for the community. Moreover, funders like the Dutch Research Council (NWO) already offer grants for replication studies.

In addition, universities and other educational institutions can be supported to adapt curricula that prioritize transparent and robust science, using resources such as FORRT's Lesson Plans (Pownall et al., 2024; https://forrt.org/lesson-plans/), Clusters (https://forrt.org/clusters/), and Curated Resources (https://forrt.org/resources/).

Lastly, researchers, science communicators, and journalists should shift away from highlighting sensational findings (see Sumner et al., 2016) and instead promote research focused on replicability, metascience, and robustness.

Judging replication results

Replication plays a critical role in ensuring the robustness and reliability of scientific claims, but it is vital to acknowledge the complexity behind failed replications.

Replication failures can arise for many reasons, and understanding these reasons is essential to fostering a constructive — rather than punitive — approach to scientific progress. Potential explanations for low replicability range from questionable research practices and publication bias to more inherent issues such as measurement error, the inherent heterogeneity of social and psychological phenomena, or the heterogeneous methods of measuring them.

One significant factor is the historic and widespread issue of low statistical power. Studies with insufficient sample sizes for the effects they are intending to examine, particularly in the social sciences, are more prone to false positives and inflated effect sizes. Furthermore, the so-called 'crud factor' — the tendency for almost everything to be weakly correlated — makes it challenging to distinguish meaningful effects from statistical noise (see Bakan, 1966; Mehl, 1990; Orben & Lakens, 2024). This means that studies with large sample sizes may detect effects that lack real-world significance, highlighting the need for new and context-dependent thresholds for clinical or practical meaningfulness when interpreting effect sizes.

Moreover, it is essential to recognize the broader context in replication outcomes. Social, cognitive, and behavioral effects are not universal and may vary across factors such as time, population, geography. Heterogeneity in study conditions can cause genuine effects to fail under different circumstances, but this does not necessarily invalidate the original findings. There is also ongoing discussion on the exact distinction between direct and conceptual replications, which may influence interpretation of replication results (see Derksen & Morawski, 2022). Instead, replication failures can help identify the above-mentioned boundary conditions, clarify where and when certain effects are likely to hold, thus adding to the diverse knowledge surrounding a certain effect. As such, it is important to approach replication failures with nuance, recognizing that they may reveal the presence of moderators or mediators rather than indicating a lack of support for the hypothesis.

The credibility revolution underscores systemic factors pushing science toward greater transparency, robustness, and replicability (Nosek et al., 2022). The increased scrutiny of research practices, the growing emphasis on open data, analysis code, and materials, and the demand for higher methodological standards, all contribute to a more accountable and reliable scientific process (see the *Registered Reports* format; Soderberg et al., 2021). While there is no consensus on how to classify replications as being on a spectrum between successful and failed, ongoing efforts focus on identifying factors that enhance replicability by analyzing replication outcomes (Boyce et al., 2023). Rather than only being seen as a failure, the credibility revolution gives us the chance to drive reform, fostering a culture of reproducibility and rigorous evaluation that strengthens the foundations of empirical research going forward (Korbmacher et al., 2023).

Replications in the future

We propose four key features that a scientific ecosystem can adopt to take full advantage of replication research going forward: 1) findability of replications, 2) widespread adoption of open science practices, 3) education and training surrounding replications, and 4) incentivizing replications.

Figure 3. The FORRT Replication Hub. The FORRT tower icon indicates that a resource is available in the FORRT Replication Hub. All other projects are currently in development.

First, replication studies should be easy to find so they can be easily incorporated in one's new research. It would be ideal if search engines could automatically tag replication studies. However, this might be an error-prone process with artificial intelligence being needed to identify which journal articles that cite earlier studies are replications of those studies and which (merely) cite the earlier studies (Google Scholar, for instance, already tracks citations). Nonetheless, human, crowd-sourced validation is likely to remain essential to guarantee accuracy and interpretative nuance, even as automated artificial intelligence systems improve, an approach we have adopted in developing the FORRT Replication Hub (see Figure 3). Our Hub consolidates human-generated replication-related projects – in particular, the Replications & Reversals project, the Replication Database, and a handbook for conducting replication studies. In addition, this hub includes a dedicated journal *Replication Research* (http://replicationresearch.org) that we are developing with stakeholder engagement. Ultimately, we envision this hub to evolve into the go-to platform for academics and students to search for and publish replication studies, to engage in interdisciplinary dialogue about replication, and to consult in methods and statistics courses. Other innovations include *PubPeer* (https://pubpeer.com/), which allows people to comment on original studies, highlight replications, and discuss conflicting results, and tools like Zotero plug-ins and Scite.ai that can flag articles with replication discussions and retraction notices, enabling more efficient literature reviews. To further address findability, we propose establishing dedicated platforms for replication studies, such as curated pre-print collections, databases, and journals dedicated to replication studies. These platforms would provide easier pathways to publish, make replication efforts findable through search functions, promote citation of replications, and enable interdisciplinary discussions on replication standards. If not published alongside each other, replication attempts could be systematically linked with the original studies to increase visibility and support cumulative research. Our Replication Hub and Database are pioneers in this process.

Second, primary research needs to adopt open science practices across the board. At a minimum, published studies should include detailed methods descriptions, open materials, open data (when ethically appropriate) and open analysis code. Moreover, research should be preregistered (if possible) or, better yet, take the form of *Registered Reports*, to ensure that confirmatory and exploratory analyses are clearly labeled. Unfortunately, transparency is still uncommon (Hardwicke et al., 2024). Furthermore, authors are not very responsive to requests to provide data (out of 65 contacted researchers from studies where data was 'available upon request', only 27% actually shared data; Hussey, 2025), and rates of data sharing differ by discipline, with psychology on the lower end (Tedersoo et al., 2021). These factors can make precise replication difficult – or impossible. For example, Errington et al. (2021) could only replicate 50 experiments from 23 papers (even though they initially set out to replicate 193 experiments from 53 high-impact papers). We suggest that journals have a key role to play: they should enforce a rule that transparency is the default for submitted manuscripts, as several journals have done recently (Hardwicke & Vazire, 2024; Wildman et al., 2024).

Third, researchers should be trained in replication-related methodologies, such as equivalence testing (Lakens, 2017, 2022), verification of original studies (Feldman, 2025), reproducibility tests (Lindsay, 2023), sample size planning and power analyses (Simonsohn, 2015), effect size and confidence intervals calculations of original studies (Jané et al., 2024), preregistrations (Brandt et al., 2014), and replication success criteria (LeBel et al., 2019). In fact, we want to stress that teaching about replication research needs to be a major cornerstone of teaching science and the scientific method (Bauer et al., 2025; Boyce et al., 2023; Frank & Saxe, 2012; Hawkins et al., 2018), for example as part of undergraduate training (Button, 2018; Chopik et al., 2018; Grahe et al., 2012; Jekel et al., 2020; Pownall et al., 2023, 2024; Wagge et al., 2019).

Fourth, replication research needs to be rewarded. Universities and funders should officially recognize the value of replication studies, particularly when they contribute new theoretical insights, methodological advancements, or extensions of prior findings, rather than sideline them, as several key funding bodies currently do. For example, the Research Excellence Framework (REF) exercise in the United Kingdom evaluates a nominated study's originality, significance, and rigour. As replications arguably score highly on significance and rigour, such evaluation exercises should make reviewers aware to provide explicit recognition and reward for replication attempts. Updating journal submission guidelines to actively promote the submission of replication studies could include a Pottery Barn rule — "you break it, you buy it" (Srivastava, 2012) — which requires journals to publish replications of studies they previously published (a policy implemented by *Royal Society Open Science*). Editors and journals might even actively suggest replication attempts for studies on which new research is building. A significant proportion of recent replication research has involved large-scale efforts with up to a hundred independent studies (see Open Science Collaboration, 2015), which prestigious journals often favor due to their extensive sample sizes and scope — something single-lab teams can rarely achieve (but see Boyce et al., 2023).

Some progress has been made, with 128 journals implementing policies that support replication studies, adhering to Transparency and Openness Promotion (TOP) Factor level 3 standards (as of June 2024). Likewise, funders such as the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) tend to support large batches of replications, which puts smaller replication studies at a disadvantage. This model is not sustainable, as replication studies in specialized areas often require specific resources, equipment, or expertise, making large-scale replications impractical for specific fields. It can take years to conduct such extensive replications, and a single replication of a longitudinal study may demand as much effort and resources as several cross-sectional replications. Given the central role of replications, a more systematic evaluation process based on cost-benefit analyses could help determine which studies are most urgently needing replication, ensuring that resources are directed where they are most valuable (Feldman, 2025; Isager et al., 2023, 2024). How to choose such studies and the metrics to evaluate them on is a key topic for future investigations.

Conclusion

Replications are intricate and complex. We recommend that the scientific community adopts a pluralistic and dynamic approach to replication — one that appreciates the various reasons why effects may fail to replicate and avoids treating every replication failure as a definitive refutation. Replications should be valued for their role in refining theories and improving the cumulative understanding of scientific phenomena. Furthermore, as we integrate replication more deeply into research practices, we have the power to ensure that these efforts are properly found, used, taught, published, and valued. Initiatives such as the FORRT Replication Hub provide a platform to make replications more visible, accessible, rewarding, and integral to scientific discourse. By systematically linking replication attempts to original studies and fostering interdisciplinary discussions and by publishing and thereby rewarding high-quality replication studies, our Replication Hub hopes to overcome the barriers that have historically limited the role of replication in science. Ultimately, replications should not be seen as a final verdict but as a dynamic part of the scientific process that drives progress through a continuous and cumulative reassessment of claims and evidence.

References

Alleva, L. (2006). Taking time to savour the rewards of slow science. *Nature, 443*, 271. https://doi.org/10.1038/443271e

Ankel-Peters, J., Fiala, N., & Neubauer, F. (2023). Do economists replicate? *Journal of Economic Behavior & Organization, 212*, 219–232. https://doi.org/10.1016/j.jebo.2023.05.009

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology, 3*(3), 266-286. https://doi.org/10.1080/23743603.2019.1684822

Anvari, F., Kievit, R., Lakens, D., Pennington, C. R., Przybylski, A. K., Tiokhin, L., Wiernik, B. M., & Orben, A. (2023). Not all effects are indispensable: Psychological science requires verifiable lines of reasoning for whether an effect matters. *Perspectives on Psychological Science, 18*(2), 503-507. https://doi.org/10.1177/17456916221091565

Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66*(6), 423. https://psycnet.apa.org/doi/10.1037/h0020412

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature, 553*, 452–454. https://doi.org/10.1038/533452a

Bauer, G., Breznau, N., Gereke, J., Höffler, J. H., Janz, N., Rahal, R., Rennstich, J.K., & Soiné, H. (2025). Teaching constructive replications in the behavioral and social sciences using quantitative data. *Teaching of Psychology, 52*(1), 117-123. https://doi.org/10.1177/00986283231219503

Bauernfeind, R. H. (1968). The need for replication in educational research. *The Phi Delta Kappan, 50*(2), 126-128. https://www.jstor.org/stable/10.5951/jresematheduc.49.1.0002

Begley, C., & Ellis, L. (2012). Raise standards for preclinical cancer research. *Nature, 483*, 531–533. https://doi.org/10.1038/483531a

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... & Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*(1), 6-10. https://doi.org/10.1038/s41562-017-0189-z

Bishop, D. V. M. (2018). Fallibility in science: Responding to errors in the work of oneself and others. *Advances in Methods and Practices in Psychological Science, 1*(3), 432-438. https://doi.org/10.1177/2515245918776632

Boekel, W., Wagenmakers, E. J., Belay, L., Verhagen, J., Brown, S., & Forstmann, B. U. (2015). A purely confirmatory replication study of structural brain-behavior correlations. *Cortex, 66*, 115-133. https://doi.org/10.1016/j.cortex.2014.11.019

Boyce, V., Mathur, M., & Frank, M. C. (2023). Eleven years of student replication projects provide evidence on the correlates of replicability in psychology. *Royal Society Open Science, 10*(11), 231240. https://doi.org/10.1098/rsos.231240

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology, 50*, 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Brodeur, A., Mikola, D., Cook, N., Brailey, T., Briggs, R., de Gendre, A. Dupraz, Y., Fiala, L., Gabani, J., Gauriot, R., Haddad, J., McWay, R., Levin, J., Johannesson, M., Metson, L., Kinge, J.M., Tian, W., Wochner, T., Mishra, S.,… & McManus, E. (2024). *Mass Reproducibility and Replicability: A New Hope. I4R Discussion Paper Series 107. The Institute for Replication (I4R).* https://hdl.handle.net/10419/289437

Burgard, T., Bosnjak, M., & Studtrucker, R. (2022). PsychOpen CAMA: Publication of community‐augmented meta‐analyses in psychology. *Research Synthesis Methods, 13*(1), 134-143. https://doi.org/10.1002/jrsm.1536

Button, K. (2018). Reboot undergraduate courses for reproducibility. *Nature, 561*(7723), 287-288. https://doi.org/10.1038/d41586-018-06692-8

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Gideon, N., Pfeiffer, T., Razan, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science, 351*(6280), 1433-1436. https://doi.org/10.1126/science.aaf0918

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B.A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T.,... & Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2*(9), 637-644. https://doi.org/10.1038/s41562-018-0399-z

Chawla, D. S. (2024). No more hunting for replication studies: crowdsourced database makes them easy to find. *Nature Index.* https://doi.org/10.1038/d41586-024-02598-w

Chivers, T. (2019). What's next for psychology's embattled field of social priming. *Nature, 576*(7786), 200-202. https://doi.org/10.1038/d41586-019-03755-2

Chopik, W. J., Bremner, R. H., Defever, A. M., & Keller, V. N. (2018). How (and whether) to teach undergraduates about the replication crisis in psychological science. *Teaching of Psychology, 45*(2), 158–163. https://doi.org/10.1177/0098628318762900

Clarke, B., Lee, P. Y., Schiavone, S. R., Rhemtulla, M., & Vazire, S. (2023). The Prevalence of Direct Replication Articles in Top-Ranking Psychology Journals. Advances in *American Psychologist.* https://doi.org/10.1037/amp0001385

Derksen, M., & Morawski, J. (2022). Kinds of replication: Examining the meanings of “conceptual replication” and “direct replication”. *Perspectives on Psychological Science, 17*(5), 1490-1505. https://doi.org/10.1177/17456916211041116

Edlund, J. E., Cuccolo, K., Irgens, M., Wagge, J.R., & Zlokovich, M.S. (2021). Saving science through replication studies. *Perspectives on Psychological Science, 16*(4), 725-743. https://doi.org/10.1177/1745691620984385

Errington, T. M., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021a). Reproducibility in Cancer Biology: Challenges for assessing replicability in preclinical cancer biology. *Elife, 10*. https://doi.org/10.7554/eLife.67995

Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021b). Investigating the replicability of preclinical cancer biology. *Elife, 10*, e71601. https://doi.org/10.7554/eLife.71601

Evans, T. R., Pownall, M., Collins, E., Henderson, E. L., Pickering, J. S., O'Mahony, A., Zaneva, M., Jaquiery, M., & Dumbalska, T. (2022). A network of change: united action on research integrity. *BMC research notes, 15*(1), 141. https://doi.org/10.1186/s13104-022-06026-y

Feldman, G. (2025). The value of replications goes beyond replicability and is associated with the value of the research it replicates: Commentary on Isager et al. (2024). *Meta-Psychology, 9*. https://doi.org/10.15626/MP.2024.4326

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science, 7*(6), 555-561. https://doi.org/10.1177/1745691612459059

Ferracioli-Oda, E., Qawasmi, A., & Bloch, M. H. (2013). Meta-analysis: melatonin for the treatment of primary sleep disorders. *PLOS ONE, 8*(5), e63773. https://doi.org/10.1371/journal.pone.0063773

Field, S. M., Hoekstra, R., Bringmann, L., & van Ravenzwaaij, D. (2019). When and why to replicate: As easy as 1, 2, 3?. *Collabra: Psychology, 5*(1), 46. https://doi.org/10.1525/collabra.218

Forscher, P. S., Wagenmakers, E. J., Coles, N. A., Silan, M. A., Dutra, N., Basnight-Brown, D., & IJzerman, H. (2023). The benefits, barriers, and risks of big-team science. *Perspectives on Psychological Science, 18*(3), 607-623. https://doi.org/10.1177/17456916221082970

Frank, M. C., & Saxe, R. (2012). Teaching replication. *Perspectives on Psychological Science, 7*(6), 600-604. https://doi.org/10.1177/1745691612460686

Frith, U. (2020). Fast lane to slow science. *Trends in Cognitive Sciences, 24*(1), 1-2. https://doi.org/10.1016/j.tics.2019.10.007

Ghai, S., Forscher, P. S., & Chuan-Peng, H. (2024). Big-team science does not guarantee generalizability. *Nature Human Behaviour, 8*(6), 1053–1056. https://doi.org/10.1038/s41562-024-01902-y

Mac Giolla, E., Karlsson, S., Neequaye, D. A., & Bergquist, M. (2022). Evaluating the replicability of social priming studies. *Preprint.* https://doi.org/10.15626/MP.2022.3308

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*(10), 535–549. https://doi.org/10.1111/spc3.12267

Grahe, J. E., Reifman, A., Hermann, A. D., Walker, M., Oleson, K. C., Nario-Redmond, M., & Wiebe, R. P. (2012). Harnessing the undiscovered resource of student research projects. *Perspectives on Psychological Science, 7*(6), 605–607. https://doi.org/10.1177/1745691612459057

Hardwicke, T. E., Thibault, R. T., Clarke, B., Moodie, N., Crüwell, S., Schiavone, S. R., ... & Vazire, S. (2024). Prevalence of transparent research practices in psychology: A cross-sectional study of empirical articles published in 2022. *Advances in Methods and Practices in Psychological Science, 7*(4), 25152459241283477.

Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., & Ioannidis, J. P. A. (2021). Estimating the Prevalence of Transparency and Reproducibility-Related Research Practices in Psychology (2014–2017). *Perspectives on Psychological Science, 17*(1), 174569162097980. https://doi.org/10.1177/1745691620979806

Hardwicke, T. E., & Vazire, S. (2024). Transparency is now the default at Psychological Science. *Psychological Science, 35*(7), 708-711. https://doi.org/10.1177/09567976231221573

Hawkins, R. X., Smith, E. N., Au, C., Arias, J. M., Catapano, R., Hermann, E., Keil, M., Lampinen, A., Raposo, S., Reynolds, J., Salehi, S., Salloum, J., Tan, J., & Frank, M. C. (2018). Improving the replicability of psychological science through pedagogy. *Advances in Methods and Practices in Psychological Science, 1*(1), 7-18. https://doi.org/10.1177/2515245917740427

Hedges, L. V., & Schauer, J. M. (2019). More than one replication study is needed for unambiguous tests of replication. *Journal of Educational and Behavioral Statistics, 44*(5), 543-570. https://doi.org/10.3102/1076998619852953

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

Hope, D., Dewar, A., & Hay, C. (2021). Is there a replication crisis in medical education research?. *Academic Medicine, 96*(7), 958-963. https://doi.org/10.1097/ACM.0000000000004063

Hüffmeier, J., Mazei, J., & Schultze, T. (2016). Reconceptualizing replication as a sequence of different studies: A replication typology. *Journal of Experimental Social Psychology, 66*, 81-92. https://doi.org/10.1016/j.jesp.2015.09.009

Hunt, K. (1975). Do we really need more replications?. *Psychological Reports, 36*(2), 587-593. https://doi.org/10.2466/pr0.1975.36.2.587

Hunter, J. E. (2001). The Desperate Need for Replications. *Journal of Consumer Research, 28*(1), 149-158. https://doi.org/10.1086/321953

Hussey, I. (2025). Data is not available upon request. *Meta-Psychology, 9.* https://doi.org/10.15626/MP.2023.4008

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Isager, P. M., van Aert, R. C. M., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., Krueger, J. I., Perugini, M., Ropovik, I., van 't Veer, A. E., Vranka, M., & Lakens, D. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods, 28*(2), 438–451. https://doi.org/10.1037/met0000438

Isager, P. M., van 't Veer, A. E., & Lakens, D. (2024). Replication value as a function of citation impact and sample size. Forthcoming in *Metapsychology*. https://doi.org/10.31222/osf.io/knjea

Jané, M. B., Xiao, Q., Yeung, S. K., Ben-Shachar, M. S., Caldwell, A. R., Cousineau, D., Dunleavy, D. J., Elsherif, M., Johnson, B. T., Moreau, D., Riesthuis, P., Röseler, L., Steele, J., Vieira, F. F., Zloteanu, M., & Feldman, G. (2024). Guide to effect sizes and confidence intervals. *PsyArXiv*. https://doi.org/10.17605/OSF.IO/D8C4G

Jekel, M., Fiedler, S., Allstadt Torras, R., Mischkowski, D., Dorrough, A. R., & Glöckner,

A. (2019). How to teach open science principles in the undergraduate curriculum—The Hagen Cumulative Science Project. *Psychology Learning & Teaching, 19*(1), 91–106. https://doi.org/10.1177/1475725719868149

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524-532. https://doi.org/10.1177/0956797611430953

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux.

Kahneman, D. (2012). A proposal to deal with questions about priming effects. *Nature News and Comments.* https://www.nature.com/news/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf

Kamermans, K. L., Dudda, L., Daikoku, T., & Verheyen, S. (2024). The is-ought problem in deciding what to replicate: which motives guide current replication practices? *PsyArXiv*. https://doi.org/10.31234/osf.io/6xdy2

Kelly, C. (2006). Replicating Empirical Research In Behavioral Ecology: How And Why It Should Be Done But Rarely Ever Is. *The Quarterly Review of Biology, 81*(3), 221–236. https://doi.org/10.1086/506236

Klein, R. A., Ratliff, K., Vianello, M., Adams Jr, R. B., Bahník, S., & Bernstein, M. J. (2014). Investigating variation in replicability: a “many labs” replication project. *Social Psychology, 45*, 142–52. https://doi.org/10.1027/1864-9335/a000178

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams Jr, R. B., Alper, S., ... & Sowden, W. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443-490. https://doi.org/10.1177/2515245918810225

Kobrock, K., & Roettger, T. B. (2023). Assessing the replication landscape in experimental linguistics. *Glossa Psycholinguistics, 2*(1). https://doi.org/10.5070/g6011135

Kohrs, F. E., Auer, S., Bannach-Brown, A., Fiedler, S., Haven, T. L., Heise, V., ... & Weissgerber, T. L. (2023). Eleven strategies for making reproducible research and open science training the norm at research institutions. *Elife, 12*, e89736. https://doi.org/10.7554/eLife.89736

Korbmacher, M., Azevedo, F., Pennington, C.R., Hartmann, H., Pownall, M., Schmidt, K.,...Evans, T.R. (2023). The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology, 1*, 3.https://doi.org/10.1038/s44271-023-00003-2

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lakens, D. (2022). Sample size justification. Collabra: *Psychology, 8*(1), 33267. https://doi.org/10.1525/collabra.33267

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science, 1*(3), 389-402. https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology, 3*. https://doi.org/10.15626/MP.2018.843

Leys, R. (2024). *Anatomy of a train wreck: The rise and fall of priming research.*

University of Chicago Press.
Loncar, T. (2021). A decade of 'power posing': Where do we stand? Tom Loncar on credibility and challenges. *Psychologist, 34*, 40-45. https://www.bps.org.uk/psychologist/decade-power-posing-where-do-we-stand
Lindsay, D. S. (2023). A plea to psychology professional societies that publish journals: Assess computational reproducibility. *Meta-Psychology, 7*. https://doi.org/10.15626/MP.2023.4020
Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher, 43*(6), 304–316. https://doi.org/10.3102/0013189X14545513
Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B., & Coyne, M. (2016). Replication of special education research: Necessary but far too rare. *Remedial and Special Education, 37*(4), 205–212. https://doi.org/10.1177/0741932516646083
Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur?. *Perspectives on Psychological Science, 7*(6), 537-542. https://doi.org/10.1177/1745691612460688
Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field: Replication in second language research. *Language Learning, 68*(2), 321–391. https://doi.org/10.1111/la ng.12286
Martin, G. N., & Clarke, R. M. (2017). Are psychology journals anti-replication? A snapshot of editorial practices. *Frontiers in Psychology, 8*, 523. https://doi.org/10.3389/fpsyg.2017.00523
McNeeley, S., & Warner, J. J. (2015). Replication in criminology: Necessary practice. *European Journal of Criminology, 12*(5), 581–597. https://doi.org/10.1177/1477370815578197
Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*(1), 195–244. https://doi.org/10.2466/PR0.66.1.195-244
Mueller-Langer, F., Fecher, B., Harhoff, D., & Wagner, G. G. (2019). Replication studies in economics—How many and which papers are chosen for replication, and why? *Research Policy, 48(*1), 62–83. https://doi.org/10.1016/j.respol.2018.07.019
Munafò, M. R., Chambers, C., Collins, A., Fortunato, L., & Macleod, M. (2022). The reproducibility debate is an opportunity, not a crisis. *BMC Research Notes, 15*(1), 43. https://doi.org/10.1186/s13104-022-05942-3
Nagy, T., Hergert, J., Elsherif, M. M., Wallrich, L., Schmidt, K., Waltzer, T., Payne, J.W., Gjoneska, B., Seetahul, Y., Wang, Y.A., Scharfenberg, D., Tyson, G., Yang, Y-F., Skvortsova, A., Alarie, S., Graves, K.A., Sotola, L.K., Moreau, D., & Rubínová, E. (2025). Bestiary of questionable research practices in psychology. *PsyArXiV*. https://doi.org/10.31234/osf.io/fhk98_v2
Nikolakopoulou, A., Mavridis, D., Furukawa, T. A., Cipriani, A., Tricco, A. C., Straus, S. E., ... & Salanti, G. (2018). Living network meta-analysis compared with pairwise meta-analysis in comparative effectiveness research: empirical study. *BMJ*, 360. https://doi.org/10.1136/bmj.k585
Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLoS Biology, 18*(3), e3000691, https://doi.org/10.1371/journal.pbio.3000691.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Struhl, M.K., Nuijten, M.B., Rohrer, J.M., Romero, F., Scheel, A.M., Scherer, L.D., Schönbrodt, F.D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73*(1), 719-748. https://doi.org/10.1146/annurev-psych-020821-114157

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). https://doi.org/10.1126/science.aac4716

Orben, A., & Lakens, D. (2020). Crud (re) defined. *Advances in Methods and Practices in Psychological Science, 3*(2), 238-247. https://doi.org/10.1177/2515245920917961

Owens, B. (2013). Long-term research: Slow science. *Nature, 495*, 300–303. https://doi.org/10.1038/495300a

Papies, E. K. (2016). Health goal priming as a situated intervention tool: How to benefit from nonconscious motivational routes to health behaviour. *Health Psychology Review, 10*(4), 408–424. https://doi.org/10.1080/17437199.2016.1183506

Parsons, S., Azevedo, F., Elsherif, M. M., Guay, S., Shahim, O. N., Govaart, G. H., Norris, E., O'Mahony, A., Parker, A. J., Todorovic, A., Pennington, C. R., Garcia-Pelegrin, E., Lazić, A., Robertson, O., Middleton, S. L., Valentini, B., McCuaig, J., Baker, B. J., Collins, E.,... & Aczel, B. (2022). A community-sourced glossary of open scholarship terms. *Nature Human Behaviour, 6*(3), 312-318. https://doi.org/10.1038/s41562-021-01269-4

Patil, P., Peng, R. D., & Leek, J. T. (2016). What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. *Perspectives on Psychological Science, 11*(4), 539-544. https://doi.org/10.1177/1745691616646366

Peetz, H. K., Primbs, M., Dudda, L., Andresen, P. K., Pennington, C. R., Westwood, S. J., & Buchanan, E. M. (2024). Evaluating interventions: A practical primer for specifying the smallest effect size of interest. *PsyArXiv.* https://doi.org/10.31234/osf.io/3qmj4

Pennington, C. (2024). Children of the replication crisis. *The Psychologist.* https://www.bps.org.uk/psychologist/children-replication-crisis

Pownall, M., Azevedo, F., Aldoh, A., Elsherif, M., Vasilev, M., Pennington, C. R., Robertson, O., Tromp, M. V., Liu, M., Makel, M. C., Tonge, N., Moreau, D., Horry, R., Shaw, J., Tzavella, L., McGarrigle, R., Talbot, C., Parsons, S., & FORRT. (2024). Embedding open and reproducible science into teaching: A bank of lesson plans and resources. *Scholarship of Teaching and Learning in Psychology, 10*(3), 342–349. https://doi.org/10.1037/stl0000307

Pownall, M., Azevedo, F., König, L. M., Slack, H. R., Evans, T. R., Flack, Z., Grinschgl, S., Elsherif, M. M., Gilligan-Lee, K. A., de Oliveira, C. M. F., Gjoneska, B., Kalandadze, T., Button, K., Ashcroft-Jones, S., Terry, J., Albayrak-Aydemir, N., Děchtěrenko, F., Alzahawi, S., & Baker, B. J., … FORRT. (2023). Teaching open and reproducible scholarship: A critical review of the evidence base for current pedagogical methods and their outcomes. *Royal Society Open Science, 10*(5), 1–28. https://doi.org/10.1098/rsos.221255

Quintana, D. S. (2021). Replication studies for undergraduate theses to improve science and education. *Nature Human Behaviour, 5*(9), 1117-1118. https://doi.org/10.1038/s41562-021-01192-8

Röseler, L., Kaiser, L., Doetsch, C., Klett, N., Seida, C., Schütz, A., Aczel, B., Adelina, N., Agostini, V., Alarie, S., Albayrak-Aydemir, N., Aldoh, A., Al-Hoorie, A. H., Azevedo, F., Baker, B. J., Barth, C. L., Beitner, J., Brick, C., Brohmer, H., ... & Zhang, Y. (2024). The replication database: Documenting the replicability of psychological science. *Journal of Open Psychology Data, 12*(8), 1-23. https://doi.org/10.5334/jopd.101

Röseler, L., Wallrich, L., Hartmann, H., & Azevedo, F. (2025). *FReD: Interfaces to the FORRT Replication Database.* R package version 0.0.0.9001, http://forrt.org/FReD/

Rosenthal, R. (1979). The "file drawer problem" and tolerance of null results. *Psychological Bulletin, 86*, 638-641. https://doi.org/10.1037/0033-2909.86.3.638

Serra-Garcia, M., & Gneezy, U. (2021). Nonreplicable publications are cited more than replicable ones. *Science Advances, 7*(21), eabd1705. https://doi.org/10.1126/sciadv.abd1705

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. Frontiers in Psychology, 10, 813. https://doi.org/10.3389/fpsyg.2019.00813

Scheel, A. M., Schijen, M. R., & Lakens, D. (2021). An excess of positive results: Comparing the standard psychology literature with registered reports. *Advances in Methods and Practices in Psychological Science, 4*(2), 25152459211007467. https://doi.org/10.1177/25152459211007467

Schimmack., U., Heene, M., & Kesavan, K. (2017). *Reconstruction of a Train Wreck: How Priming Research Went off the Rails.* https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-of-the-rails/

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

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

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26*(5), 559–569. https://doi.org/10.1177/0956797614567341

Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J., Thorn, F. S., Vazire, S., ... & Nosek, B. A. (2021). Initial evidence of research quality of registered reports compared with the standard publishing model. *Nature Human Behaviour, 5*(8), 990-997. https://doi.org/10.1038/s41562-021-01142-4

Srivastava, S. (2012). A Pottery Barn rule for scientific journals. *The Hardest Science.* https://thehardestscience.com/2012/09/27/a-pottery-barn-rule-for-scientific-journals/

Stengers, I. (2017). *Another science is possible: A manifesto for slow science.* Cambridge, UK: Polity Press.

Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Bott, L., Adams, R., ... & Chambers, C. D. (2016). Exaggerations and caveats in press releases and health-related science news. *PloS One, 11*(12), e0168217.

https://doi.org/10.1371/journal.pone.0168217
Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., ... & Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data, 8*(1), 192. https://doi.org/10.1038/s41597-021-00981-0
Tiokhin, L., Hackman, J., Munira, S., Jesmin, K., & Hruschka, D. (2019). Generalizability is not optional: Insights from a cross-cultural study of social discounting. *Royal Society Open Science, 6*(2), 181386. https://doi.org/10.1098/rsos.181386
Tunç, D. U., & Tunç, M. N. (2023). A falsificationist treatment of auxiliary hypotheses in social and behavioral sciences: Systematic replications framework. *Meta-Psychology, 7*. https://doi.org/10.15626/MP.2021.2756
Urminsky, O., & Dietvorst, B. J. (2024). Taking the full measure: Integrating replication into research practice to assess generalizability. *Journal of Consumer Research, 51*(1), 157-168. https://doi.org/10.1093/jcr/ucae007
Wagenmakers, E.-J., Wetzels, R., Borsboom, D., and van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: Comment on Bem. *Journal of Personality and Social Psychology, 100*, 426–432. https://doi.org/10.1037/a0022790
Wagge, J. R., Brandt, M. J., Lazarevic, L. B., Legate, N., Christopherson, C., Wiggins, B., & Grahe, J. E. (2019). Publishing research with undergraduate students via replication work: The collaborative replications and education project. *Frontiers in Psychology, 10*, 247. https://doi.org/10.3389/fpsyg.2019.00247
Wildman, W. J., Ross, R. M., McKay, R. T., Watts, J., Hoogeveen, S., Sosis, R., ... & Shaver, J. (2024). Religion, Brain & Behavior adopts stricter transparency standards. *Religion, Brain & Behavior, 14*(4), 341-344. https://doi.org/10.1080/2153599X.2024.2424031
Wingen, T., Berkessel, J. B., & Englich, B. (2020). No replication, no trust? How low replicability influences trust in psychology. *Social Psychological and Personality Science, 11*(4), 454-463. https://doi.org/10.1177/1948550619877412
Yong, E (2012). Nobel laureate challenges psychologists to clean up their act. *Nature*. https://doi.org/10.1038/nature.2012.11535
Zhu, J., Cai, H., & Shoda, Y. (2025). How often are constraints on generality (COGs) discussed? A look at articles published in four Psychology journals 2018-2022. *Collabra: Psychology, 11*(1), 128652. https://doi.org/10.1525/collabra.128652
Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences, 41*, e120. https://psycnet.apa.org/doi/10.1017/S0140525X17001972