A community-sourced glossary of open scholarship terms

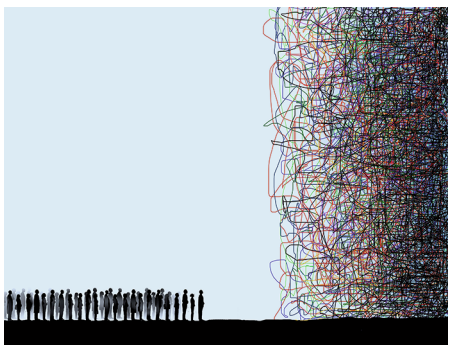
Open scholarship has transformed research, and introduced a host of new terms in the lexicon of researchers. The 'Framework for Open and Reproducible Research Teaching' (FORRT) community presents a crowdsourced glossary of open scholarship terms to facilitate education and effective communication between experts and newcomers.

Sam Parsons, Flávio Azevedo, Mahmoud M. Elsherif, Samuel Guay, Owen N. Shahim, Gisela H. Govaart, Emma Norris, Aoife O'Mahony, Adam J. Parker, Ana Todorovic, Charlotte R. Pennington, Elias Garcia-Pelegrin, Aleksandra Lazić, Olly Robertson, Sara L. Middleton, Beatrice Valentini, Joanne McCuaig, Bradley J. Baker, Elizabeth Collins, Adrien A. Fillon, Tina B. Lonsdorf, Michele C. Lim, Norbert Vanek, Marton Kovacs, Timo B. Roettger, Sonia Rishi, Jacob F. Miranda, Matt Jaquiery, Suzanne L. K. Stewart, Valeria Agostini, Andrew J. Stewart, Kamil Izydorczak, Sarah Ashcroft-Jones, Helena Hartmann, Madeleine Ingham, Yuki Yamada, Martin R. Vasilev, Filip Dechterenko, Nihan Albayrak-Aydemir, Yu-Fang Yang, Annalise A. LaPlume, Julia K. Wolska, Emma L. Henderson, Mirela Zaneva, Benjamin G. Farrar, Ross Mounce, Tamara Kalandadze, Wanyin Li, Qinyu Xiao, Robert M. Ross, Siu Kit Yeung, Meng Liu, Micah L. Vandegrift, Zoltan Kekecs, Marta K. Topor, Myriam A. Baum, Emily A. Williams, Asma A. Assaneea, Amélie Bret, Aidan G. Cashin, Nick Ballou, Tsvetomira Dumbalska, Bettina M. J. Kern, Claire R. Melia, Beatrix Arendt, Gerald H. Vineyard, Jade S. Pickering, Thomas R. Evans, Catherine Laverty, Eliza A. Woodward, David Moreau, Dominique G. Roche, Eike M. Rinke, Graham Reid, Eduardo Garcia-Garzon, Steven Verheyen, Halil E. Kocalar, Ashley R. Blake, Jamie P. Cockcroft, Leticia Micheli, Brice Beffara Bret, Zoe M. Flack, Barnabas Szaszi, Markus Weinmann, Oscar Lecuona, Birgit Schmidt, William X. Ngiam, Ana Barbosa Mendes, Shannon Francis, Brett J. Gall, Mariella Paul, Connor T. Keating, Magdalena Grose-Hodge, James E. Bartlett, Bethan J. Iley, Lisa Spitzer, Madeleine Pownall, Christopher J. Graham, Tobias Wingen, Jenny Terry, Catia Margarida F. Oliveira, Ryan A. Millager, Kerry J. Fox, Alaa AlDoh, Alexander Hart, Olmo R. van den Akker, Gilad Feldman, Dominik A. Kiersz, Christina Pomareda, Kai Krautter, Ali H. Al-Hoorie and Balazs Aczel

Open scholarship is an umbrella term that refers to the endeavour to improve openness, integrity, social justice, diversity, equity, inclusivity and accessibility in all areas of scholarly activities. Open scholarship extends the more widely used terms open science and open research to include academic fields beyond the sciences and activities beyond academic research.

Barriers to open scholarship terminology

Over the past decade, open scholarship has radically changed the way we think about and discuss research and higher education. New concepts, tools and practices have been developed and promoted, which introduce novel terms or repurpose existing ones. These changes have increased the breadth, but also the ambiguity, of terminology, which creates barriers to effective understanding and communication for novices and experts. Some terms, such as replicability or reproducibility, are well-known but frequently used interchangeably or differentially among fields and disciplines; other terms, such as CARKing, PARKing, or paradata, are less well-known beyond a small circle of researchers. Terms that become conventional within a given field often reflect the preferences of those with the platforms and privileges to determine academic discourse and, consequently, can act as a barrier to participation by those without such platforms. A similar barrier is that much academic language is contained within a 'hidden curriculum', meaning these terms and practices are often used under the misplaced assumption that students, or those unfamiliar with an area, understand them.



Credit: Fanatic Studio/Gary Waters/SCIENCE PHOTO LIBRARY / Getty

A diversity of new terminologies

Terms associated with open scholarship are diverse in many aspects. Some are neologisms (that is, newly coined), whereas others are reclamations of older terms (for example, *P* hacking and adversarial

Table 1 | Examples from the open scholarship glossary version 1.0

Term	Definition
Analytic flexibility	Analytic flexibility refers specifically to the large number of choices made during data preprocessing and statistical analysis. Analytic flexibility can be problematic as this variability in analytic strategies can translate into variability in research outcomes, particularly when several strategies are applied but not transparently reported.
#bropenscience	#bropenscience is a tongue-in-cheek expression intended to raise awareness of the lack of diverse voices in open science, in addition to the presence of behaviour and communication styles that can be toxic or exclusionary. Importantly, not all 'bros' are men; rather, they are individuals who demonstrate rigid thinking, lack self-awareness, and tend towards hostility, unkindness and exclusion. They generally belong to dominant groups who benefit from structural privileges. To address #bropenscience, researchers should examine and address structural inequalities within academic systems and institutions.
CARKing	'Critiquing after the results are known' (CARKing) refers to presenting a post hoc criticism of a design as one that would have been made in advance of the results being known. It usually forms a reaction to or criticism of unwelcome or unfavourable results, whether the critic is conscious of this fact or not.
Codebook	A codebook is a high-level summary that describes the contents, structure, nature and layout of a dataset. A well-documented codebook contains information intended to be complete and self-explanatory for each variable in a data file, such as the wording and coding of the item, and the underlying construct. It provides transparency to researchers who may be unfamiliar with the data but wish to reproduce analyses or reuse the data.
Conceptual replication	A replication attempt whereby the primary effect of interest is the same but tested in a different sample and captured in a different way to that originally reported (that is, using different operationalizations, data processing and statistical approaches and/or different constructs). The purpose of a conceptual replication is often to explore what conditions limit the extent to which an effect can be observed and generalized (for example, only within certain contexts, with certain samples or using certain measurement approaches) towards evaluating and advancing theory.
Creative destruction approach	Replication efforts should seek not only to support or question the original findings, but also to replace them with revised, stronger theories with greater explanatory power. This approach therefore involves pruning existing theories, comparing all the alternative theories and making replication efforts more generative and engaged in theory-building.
Credibility revolution	The problems and the solutions resulting from a growing distrust in scientific findings, following concerns about the credibility of scientific claims (for example, low replicability). The term has been proposed as a more positive alternative to the term replicability crisis, and includes the many solutions to improve the credibility of research, such as preregistration, transparency and replication.
CRediT	CRediT is a high-level taxonomy used to indicate the roles that are typically adopted by contributors to scientific scholarly output. There are currently 14 roles that describe each contributor's specific contribution to the scholarly output. They can be assigned multiple times to different authors and one author can also be assigned multiple roles. CRediT includes the following roles: 'Conceptualization', 'Data curation', 'Formal analysis', 'Funding acquisition', 'Investigation', 'Methodology', 'Project administration', 'Resources', 'Software', 'Supervision', 'Validation', 'Visualization', 'Writing – original draft' and 'Writing – review & editing'.
Decolonization	Coloniality can be described as the naturalization of concepts such as imperialism, capitalism and nationalism. Together, these concepts can be thought of as a matrix of power (and power relations) that can be traced to the colonial period. Decoloniality seeks to break down and decentralize those power relations, with the aim of understanding their persistence and reconstructing the norms and values of a given domain. In an academic setting, decolonization refers to the rethinking of the lens through which we teach, research and co-exist, so that the lens generalizes beyond Western-centred and colonial perspectives. Decolonizing academia involves reconstructing the historical and cultural frameworks being used, redistributing a sense of belonging in universities, and empowering and including voices and knowledge types that have historically been excluded from academia. This is done when people engage with their past, present and future while holding a perspective that is separate from the socially dominant perspective, and also by including — not rejecting — an individuals' internalized norms and taboos from the specific colony.
Direct replication	As direct replication does not have a widely agreed technical meaning and there is no clear-cut distinction between a direct and conceptual replication, below we list several contributions towards a consensus. Rather than debating the exactness of a replication, it is more helpful to discuss the relevant differences between a replication and its target, and their implications for the reliability and generality of the target's results. Generally, direct replication refers to a new data collection that attempts to replicate original studies' methods as closely as possible. In this sense, direct replication is a replication attempt that aims to duplicate the needed elements that produced the original results. The purpose of a direct replication can be to identify type 1 errors and/or experimenter effects, determine the replicability of an effect using the same or improved practices, or to create more specific estimates of effect size. Directness of replication is a continuum between repeating specific observations (data) and observing generalized effects (phenomena). How closely a replication replicates an original study is often a matter of debate, with differences often being cited as hidden moderators of effects. Furthermore, there can be debate over the relevant importance of technical equivalence (that is, using identical materials) versus psychological equivalence (that is, realizing the identical psychological conditions) to the original study. For example, consider a study on trust in the US president conducted in 2018. A technically equivalent replication would use Trump as stimulus (he was president in 2018), whereas a psychologically equivalent study would use Biden (he is the current president).
External validity	External validity refers to whether the findings of a scientific study can be generalized to other contexts outside of the study context (different measures, settings, people, places and times). Statistically, threats to external validity may reflect interactions whereby the effect of one factor (the independent variable) depends on another factor (a confounding variable). External validity may also be limited by the study design (for example, an artificial laboratory setting or a non-representative sample). An alternative definition for external validity, in psychometrics, explains it as the degree of evidence that confirms the relations of a tested psychological construct with external variables.

Continued

Table 1 | Examples from the open scholarship glossary version 1.0 (Continued)

Term	Definition
FAIR principles	These principles describe making scholarly materials findable, accessible, interoperable and reusable (FAIR). Findable and accessible are terms concerned with where materials are stored (for example, in data repositories), while the interoperable and reusable terms focus on the importance of data formats and how such formats might change in the future.
Garden of forking paths	This term refers to the typically invisible decision tree traversed during operationalization and statistical analysis given that "there is a one-to-many mapping from scientific to statistical hypotheses"[2]. In other words, even in absence of *P* hacking or fishing expeditions and when the research hypothesis was posited ahead of time, there can be a plethora of statistical results that can appear to be supported by theory given data. "The problem is there can be a large number of *potential* comparisons when the details of data analysis are highly contingent on data, without the researcher having to perform any conscious procedure of fishing or examining multiple p-values"[2]. The term aims to highlight the uncertainty ensuing from idiosyncratic analytical and statistical choices in mapping theory-to-test, and contrasting intentional (and unethical) questionable research practices (for example, *P* hacking and fishing expeditions) with nonintentional research practices that can, potentially, have the same effect despite not having intent to corrupt their results. The garden of forking paths refers to the decisions during the scientific process that inflate the false-positive rate as a consequence of the potential paths that could have been taken (had other decisions been made).
HARKing	This term denotes a questionable research practice: 'hypothesizing after the results are known' (HARKing). HARKing has been defined as a post hoc hypothesis that is based on or informed by a result in a research report as if it was, in fact, a priori: for example, performing subgroup analyses, finding an effect in one subgroup and writing the introduction with a 'hypothesis' that matches these results.
Incentive structure	This term denotes the set of evaluation and reward mechanisms (explicit and implicit) for scientists and their work. Incentivized areas within the broader structure include hiring and promotion practices, track record for awarding funding, and prestige indicators such as publication in journals with high impact factors, invited presentations, editorships and awards. It is commonly believed that these criteria are often misaligned with the ethos of science, and therefore do not promote rigorous scientific output. Initiatives such as the Declaration on Research Assessment (DORA)aim to reduce the field's dependency on evaluation criteria such as journal impact factors in favour of assessments based on the intrinsic quality of research outputs.
Inclusion	Inclusion or inclusivity refers to a sense of welcome and respect within a given collaborative project or environment (such as academia); diversity simply indicates a wide range of backgrounds, perspectives and experiences, whereas efforts to increase inclusion go further to promote engagement and equal valuation among diverse individuals, who might otherwise be marginalized. Increasing inclusivity often involves minimizing the effect of, or even removing, systemic barriers to accessibility and engagement.
Metadata	Metadata are structured data that describe and synthesize other data. Metadata can help to find, organize and understand data. Examples of metadata include creator, title, contributors, keywords and tags, as well as any kind of information necessary to verify and understand the results and conclusions of a study, such as codebook on data labels, descriptions, and the sample and data collection process. An alternative definition for this term is 'data about data'.
Multi-analyst studies	In typical empirical studies, a single researcher or research team conducts the analysis, which creates uncertainty about the extent to which the choice of analysis influences the results. In multi-analyst studies, two or more researchers independently analyse the same research question or hypothesis on the same dataset. A multi-analyst approach may be beneficial in increasing our confidence in a particular finding; uncovering the effect of analytical preferences across research teams; and highlighting the variability in such analytical approaches.
Open scholarship	Open scholarship is often used synonymously with open science, but extends to all disciplines, drawing in those which might not traditionally identify as science-based. It reflects the idea that knowledge of all kinds should be openly shared, transparent, rigorous, reproducible, replicable, accumulative and inclusive (allowing for all knowledge systems). Open scholarship is not solely limited to research, and includes all scholarly activities (such as teaching and pedagogy).
Open science	This is an umbrella term that reflects the idea that scientific knowledge of all kinds, where appropriate, should be openly accessible, transparent, rigorous, reproducible, replicable, accumulative and inclusive, all which are considered fundamental features of the scientific endeavour. Open science consists of principles and behaviours that promote transparent, credible, reproducible and accessible science. Open science has six major aspects: open data, open methodology, open source, open access, open peer review and open educational resources.
ORCID iD	This term refers to an organization that provides a registry of persistent unique identifiers (ORCID iDs) for researchers and scholars, allowing these users to link their digital research documents and other contributions to their ORCID record. This avoids the name ambiguity problem in scholarly communication. ORCID iDs provide unique, persistent identifiers connecting researchers and their scholarly work. It is free to register for an ORCID iD at https://orcid.org/register.
P hacking	*P* hacking is exploiting techniques that may artificially increase the likelihood of obtaining a statistically significant result by meeting the standard statistical significance criterion (typically $\alpha = 0.05$). For example, performing multiple analyses and reporting only those at $P < 0.05$, selectively removing data until $P < 0.05$ or selecting variables for use in analyses based on whether those parameters are statistically significant.
Paper mill	A paper mill is an organization that is engaged in scientific misconduct, wherein multiple papers are produced by falsifying or fabricating data; for example, by editing figures or numerical data or plagiarizing written text. A paper mill relates to the fast production and dissemination of multiple allegedly new papers. These are often not detected in the scientific publishing process and therefore either never found or retracted if discovered (for example, through plagiarism software).

Continued

Table 1 | Examples from the open scholarship glossary version 1.0 (Continued)

Term	Definition
Paradata	Paradata are data that are captured about the characteristics and context of primary data collected from an individual, and are distinct from metadata. Paradata can be used to investigate a respondent's interaction with a survey or an experiment on a micro-level. They can be most easily collected during computer-mediated surveys but are not limited to them. Examples include response times to survey questions, repeated patterns of responses such as choosing the same answer for all questions, contextual characteristics of the participant such as injuries that prevent good performance on tasks, and the number of premature responses to stimuli in an experiment. Paradata have been used for the investigation and adjustment of measurement and sampling errors.
PARKing	PARKing (preregistering after results are known) is defined as a practice in which researchers complete an experiment (possibly with infinite re-experimentation) before preregistering. This practice invalidates the purpose of preregistration, and is a questionable research practice (or even a form of scientific misconduct) that tries to gain the credibility that the study has been preregistered.
Preregistration	Preregistration is the practice of publishing the plan for a study, including research questions and hypotheses, research design and data analysis, before the data have been collected or examined. It is also possible to preregister secondary data analyses. A preregistration document is time-stamped and typically registered with an independent party (for example, a repository) so that it can be publicly shared with others (possibly after an embargo period). Preregistration provides transparent documentation of what was planned at a certain time point, and allows third parties to assess what changes may have occurred afterwards. The more detailed a preregistration is, the better third parties can assess these changes and, with that, the validity of the performed analyses. Preregistration aims to clearly distinguish confirmatory from exploratory research.
Replicability	Replicability is an umbrella term, used differently across fields, that covers concepts of: direct and conceptual replication, computational reproducibility and replicability, generalizability analysis and robustness analyses. Some of the definitions used previously include: a different team arriving at the same results using the original author's artefacts; a study arriving at the same conclusion after collecting new data; as well as studies for which any outcome would be considered diagnostic evidence about a claim from prior research.
Reproducibility	This refers to a minimum standard on a spectrum of activities ('reproducibility spectrum') for assessing the value or accuracy of scientific claims based on the original methods, data and code; for instance, where the original researcher's data and computer codes are used to re-generate the results (often referred to as computational reproducibility). Reproducibility does not guarantee the quality, correctness or validity of the published results. In some fields, this meaning is, instead, associated with the term 'replicability' or 'repeatability'.
Registered Reports	This describes a scientific publishing format that includes an initial round of peer review of the background and methods (study design, measurement and analysis plan); sufficiently high-quality manuscripts are accepted for in-principle acceptance at this stage. Typically, this stage 1 review occurs before data collection; however, secondary data analyses are possible in this publishing format. Following data analyses and write-up of results and discussion sections, the stage 2 review assesses whether authors sufficiently followed their study plan and reported deviations from it (and remains indifferent to the results). This shifts the focus of the review to the study's proposed research question and methodology and away from the perceived interest in the study's results.
Under-representation	Not all voices, perspectives, and members of the community are adequately represented. Under-representation typically occurs when the voices or perspectives of one group dominate, resulting in the marginalization of another. This often affects groups who are a minority in relation to certain personal characteristics.
WEIRD	This acronym refers to 'Western, educated, industrialized, rich and democratic' societies. Most research is conducted on and by relatively homogeneous samples from WEIRD societies. This limits the generalizability of a large number of research findings, particularly given that WEIRD people are often psychological outliers. It has been argued that 'WEIRD psychology' started to evolve culturally as a result of societal changes and religious beliefs in the Middle Ages in Europe. Critics of this term suggest it presents a binary view of the global population and erases variation that exists both between and within societies, and that other aspects of diversity are not captured.

Entries are reproduced or modified, with permission, from those in the complete glossary (all rights reserved). The complete glossary, including key references and links to related terms, can be found in Supplementary Information and is also available at https://forrt.org/glossary. This glossary is a first step in creating a common language for these concepts, facilitating discussions about the strengths and weaknesses of different open scholarship practices, and ultimately helping to build a stronger research community[8].

collaboration). Furthermore, the frequent use of acronyms can hinder immediate understanding. For example, ORCID iD[1] refers to a persistent unique identifier for individuals in their role as creator or contributor. It enables the linking of digital documents and other contributions to their digital records, attributes credit and resolves name ambiguities. Other terms use metaphors, such as the 'garden of forking paths'[2], which highlights the many alternative paths that researchers can embark on when analysing data. These examples highlight the complexity of open scholarship terminology, which often assumes prior knowledge; this assumption makes it difficult for individuals who are not versed in these terms to engage in ongoing conversations about these topics.

Communication across disciplines, and across markedly varying levels of subject and technical expertise, can be extremely difficult. Challenges can arise in understanding scientific texts when words with a historical meaning gain a new one in certain contexts. For example, the term paper mills[3] typically refers to factories that are devoted to manufacturing paper, but — in the context of open scholarship — it also denotes unethical for-profit organizations that create and publish on-demand fraudulent scientific papers that use techniques such as the fabrication of data and plagiarism.

A similar challenge arises when the meaning and usage of terms differs between (sub)disciplines and methodologies. For example, in social science fields the term preregistration refers to an uneditable, timestamped version of a research protocol, whereas in healthcare fields it refers to an accelerated course that qualifies students to fast-track into a medical profession. As another example, social scientists understand the term external validity to mean that the findings can be generalized to other contexts (different measures, people, places and times),

whereas psychometricians regard it as the relationship of a psychological concept with theoretically relevant extrinsic variables. Creative destruction[4] is yet another example: in economics, it refers to revolutionizing the economic structure from within by destroying the old system and replacing it with a new one. In psychology, a creative destruction approach to replication means that replications can — in particular circumstances — be leveraged to not only support or question the original findings, but also to replace weaker theories with stronger ones that have greater explanatory power (by preregistering different theoretical predictions and adding new measures, conditions and populations that facilitate competitive theory testing). As interdisciplinary collaboration is growing and often required by many funding agencies and stakeholders, this creates a potential for miscommunication and confusion when using such terminology.

The clarification of scholarly terminology is a challenging endeavour. It should be built on the insights of a community of experts with different perspectives and requires consensus among the members across disciplines. As the breadth of these initiatives can be overwhelming, digestible introductions to the language of open scholarship are needed[5–7]. To reduce barriers to entry and improve understanding of open scholarship terminology, as well as to foster the accessibility, inclusivity and clarity of its language, a community-driven glossary using a consensus-based methodology could help to clarify terminologies and aid in the mentoring and teaching of these concepts.

The open scholarship glossary project

The glossary presented in this Comment has been developed by the FORRT[7] community, an educational and meta-scientific initiative that aims to integrate open and reproducible research principles into higher education as well as to support educators and mentors to address related pedagogical challenges. The work has been completed in three rounds. First, the lead team created an initial list of terms related to open scholarship, and a structure for each term. Each term was required to have (1) a concise definition with supporting (2) references, (3) related terms and (4) any applicable alternative definitions. This glossary has been developed using a crowdsourced methodology with the involvement of over 100 contributors at various career stages and from a diverse range of disciplines, including psychology, economics, neuroscience, information science, social science, biology, ecology, public health and linguistics. In the second round, and in a dynamic and iterative crowdsourced process, members of the research community were invited through social media platforms or organizations such as ReproducibiliTea (https://reproducibilitea.org/) to participate in the project. The community contributors suggested new terms, to which they provided main and alternative definitions, as well as reviewed and edited other terms iteratively throughout the project. We recorded these contributions, and this is reflected in our Contributor Roles Taxonomy (CRediT; https://casrai.org/credit/) statement; all contributors were invited to be authors on this manuscript. We considered definitions as ready for dissemination when they had been reviewed by a sufficient number of contributors (typically five or more) and reached consensus. Through this process, the community-driven procedure used to develop the glossary deliberately centred the open scholarship ethos of accessibility, diversity, equity and inclusion.

The project resulted in the drafting of more than 250 terms. In Table 1, we present an abbreviated set of 30 terms (see https://forrt.org/glossary for the complete glossary, including key references and links to related terms, as well as a more detailed explanation of the project's mission and goals). The FORRT glossary project is licensed under a CC BY-NC-SA 4.0 licence; the present glossary is the 1.0 version. The version-controlled source code of the new releases of the complete glossary is archived on FORRT's website, GitHub, OSF and Zenodo, where new releases will also be stored. We set up a system allowing for continuous improvement, extension and updating from community feedback and involvement. Versioning will also allow the study of the evolution of terminologies.

As with all terminologies, this glossary will be subject to iterative improvement and updates. We encourage the scientific community to read the terms with critical eyes and to provide feedback and recommendations on FORRT's website, where instructions on how to contribute are provided and where the live version of all defined terms is publically available in its entirety. ❐

Sam Parsons[1,98], Flávio Azevedo[2,98] ✉, Mahmoud M. Elsherif[3,4,98], Samuel Guay[5], Owen N. Shahim[6], Gisela H. Govaart[7,8], Emma Norris[9], Aoife O'Mahony[10], Adam J. Parker[11], Ana Todorovic[1], Charlotte R. Pennington[12], Elias Garcia-Pelegrin[13], Aleksandra Lazić[14], Olly Robertson[15], Sara L. Middleton[16,17], Beatrice Valentini[18], Joanne McCuaig[19], Bradley J. Baker[20,21], Elizabeth Collins[22], Adrien A. Fillon[23], Tina B. Lonsdorf[24], Michele C. Lim[1], Norbert Vanek[25,26], Marton Kovacs[27,28], Timo B. Roettger[29], Sonia Rishi[30], Jacob F. Miranda[31], Matt Jaquiery[1], Suzanne L. K. Stewart[32], Valeria Agostini[30], Andrew J. Stewart[33], Kamil Izydorczak[34], Sarah Ashcroft-Jones[1], Helena Hartmann[35,36], Madeleine Ingham[30], Yuki Yamada[37], Martin R. Vasilev[38], Filip Dechterenko[39], Nihan Albayrak-Aydemir[40,41], Yu-Fang Yang[42], Annalise A. LaPlume[43], Julia K. Wolska[44], Emma L. Henderson[45], Mirela Zaneva[1], Benjamin G. Farrar[13], Ross Mounce[46], Tamara Kalandadze[47], Wanyin Li[30], Qinyu Xiao[48], Robert M. Ross[49], Siu Kit Yeung[48], Meng Liu[50], Micah L. Vandegrift[51], Zoltan Kekecs[27,52], Marta K. Topor[45], Myriam A. Baum[53], Emily A. Williams[54], Asma A. Assaneea[19,55], Amélie Bret[56], Aidan G. Cashin[57,58], Nick Ballou[59], Tsvetomira Dumbalska[1], Bettina M. J. Kern[35], Claire R. Melia[60,61], Beatrix Arendt[62], Gerald H. Vineyard[63], Jade S. Pickering[64,65], Thomas R. Evans[66], Catherine Laverty[30], Eliza A. Woodward[67], David Moreau[68], Dominique G. Roche[69], Eike M. Rinke[70,71], Graham Reid[1], Eduardo Garcia-Garzon[72], Steven Verheyen[73], Halil E. Kocalar[74], Ashley R. Blake[19], Jamie P. Cockcroft[75], Leticia Micheli[76], Brice Beffara Bret[56], Zoe M. Flack[77], Barnabas Szaszi[27], Markus Weinmann[78,79], Oscar Lecuona[80,81], Birgit Schmidt[82], William X. Ngiam[83], Ana Barbosa Mendes[84], Shannon Francis[3], Brett J. Gall[85], Mariella Paul[86], Connor T. Keating[30], Magdalena Grose-Hodge[19], James E. Bartlett[87], Bethan J. Iley[88,89], Lisa Spitzer[90], Madeleine Pownall[54], Christopher J. Graham[91], Tobias Wingen[92], Jenny Terry[93], Catia Margarida F. Oliveira[75], Ryan A. Millager[94], Kerry J. Fox[77], Alaa AlDoh[93], Alexander Hart[53], Olmo R. van den Akker[95], Gilad Feldman[48], Dominik A. Kiersz[96], Christina Pomareda[30], Kai Krautter[53], Ali H. Al-Hoorie[97] and Balazs Aczel[27]

[1]*Department of Experimental Psychology, University of Oxford, Oxford, UK.* [2]*Institute of Communication Science, Friedrich Schiller University Jena, Jena, Germany.* [3]*Department of Psychology, University of Birmingham, Birmingham, UK.* [4]*Blavatnik School of Government, University of Oxford, Oxford, UK.* [5]*Department of Psychology, University of Montreal, Montreal, Quebec, Canada.* [6]*East Midlands School of Anaesthesia, Oxford, UK.* [7]*Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.*

[8]Charité – Universitätsmedizin Berlin, Einstein Center for Neurosciences Berlin, Berlin, Germany. [9]Department of Health Sciences, Brunel University London, Uxbridge, UK. [10]School of Psychology, Cardiff University, Cardiff, UK. [11]Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, UK. [12]School of Psychology, Aston University, Birmingham, UK. [13]Department of Psychology, University of Cambridge, Cambridge, UK. [14]Department of Psychology, University of Belgrade, Belgrade, Serbia. [15]Departments of Psychiatry and Experimental Psychology, University of Oxford, Oxford, UK. [16]Department of Zoology, University of Oxford, Oxford, UK. [17]Department of Plant Sciences, University of Oxford, Oxford, UK. [18]Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland. [19]Department of English Language and Applied Linguistics, University of Birmingham, Birmingham, UK. [20]Department of Sport and Recreation Management, Temple University, Philadelphia, PA, USA. [21]Mark H. McCormack Department of Sport Management, University of Massachusetts, Amherst, MA, USA. [22]Department of Psychology, University of Stirling, Stirling, UK. [23]Department of Psychology, University of Aix-Marseille, Marseille, France. [24]Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [25]School of Cultures, Languages and Linguistics, University of Auckland, Auckland, New Zealand. [26]Experimental Research on Central European Languages Lab, Charles University Prague, Prague, Czech Republic. [27]Institute of Psychology, ELTE, Eotvos Lorand University, Budapest, Hungary. [28]Doctoral School of Psychology, ELTE Eotvos Lorand University, Budapest, Hungary. [29]Department of Linguistics and Scandinavian Studies, Universitetet i Oslo, Oslo, Norway. [30]School of Psychology, University of Birmingham, Birmingham, UK. [31]Department of Psychology, University of Alabama, Tuscaloosa, AL, USA. [32]School of Psychology, University of Chester, Chester, UK. [33]Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK. [34]Faculty of Psychology, SWPS University of Social Sciences and Humanities, Warsaw, Poland. [35]Department of Cognition, Emotion, and Methods in Psychology, University of Vienna, Vienna, Austria. [36]Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands. [37]Faculty of Arts and Science, Kyushu University, Fukuoka, Japan. [38]Department of Psychology, Bournemouth University, Bournemouth, UK. [39]Department of Mathematics, College of Polytechnics Jihlava, Jihlava, Czech Republic. [40]London School of Economics and Political Science, London, UK. [41]Open University, Milton Keynes, UK. [42]Department of Psychology, University of Würzburg, Würzburg, Germany. [43]Rotman Research Institute, Baycrest Hospital (fully affiliated with the University of Toronto), Toronto, Ontario, Canada. [44]Department of Psychology, Manchester Metropolitan University, Manchester, UK. [45]School of Psychology, University of Surrey, Guildford, UK. [46]Arcadia Fund, London, UK. [47]Department of Education, ICT and Learning, Østfold University College, Halden, Norway. [48]Department of Psychology, University of Hong Kong, Hong Kong, Hong Kong SAR. [49]Department of Psychology, Macquarie University, Sydney, New South Wales, Australia. [50]Faculty of Education, University of Cambridge, Cambridge, UK. [51]Open Knowledge Librarian, NC State University, Raleigh, NC, USA. [52]Department of Psychology, Lund University, Lund, Sweden. [53]Research Group Applied Statistical Modeling, Saarland University, Saarbrücken, Germany. [54]School of Psychology, University of Leeds, Leeds, UK. [55]Taibah University, Medina, Saudi Arabia. [56]Department of Psychology, University of Nantes, Nantes, France. [57]Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, New South Wales, Australia. [58]School of Health Sciences, University of New South Wales, Sydney, New South Wales, Australia. [59]Game AI Group, Queen Mary University of London, London, UK. [60]Department of Mental health and Social Work, Middlesex University, London, UK. [61]Department of Psychology, Keele University, Keele, UK. [62]Center for Open Science, Charlottesville, VA, USA. [63]Department of Psychology, University of East London, London, UK. [64]School of Psychology, University of Southampton, Southampton, UK. [65]Division of Neuroscience & Experimental Psychology, University of Manchester, Manchester, UK. [66]School of Human Sciences, University of Greenwich, London, UK. [67]Department of Psychology, Oxford Brookes University, Oxford, UK. [68]School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand. [69]Department of Biology, Carleton University, Ottawa, Ontario, Canada. [70]School of Politics and International Studies, University of Leeds, Leeds, UK. [71]Mannheim Centre for European Social Research (MZES), University of Mannheim, Mannheim, Germany. [72]Faculty of Health, Universidad Camilo José Cela, Madrid, Spain. [73]Department of Psychology, Education and Child Studies; Erasmus University Rotterdam, Rotterdam, The Netherlands. [74]Department of Educational Science, Muğla Sıtkı Koçman University, Muğla, Turkey. [75]Department of Psychology, University of York, York, UK. [76]Institute of Psychology, Leibniz University Hannover, Hanover, Germany. [77]School of Humanities and Social Science, University of Brighton, Brighton, UK. [78]Faculty of Management, Economics and Social Sciences, University of Cologne, Cologne, Germany. [79]Department of Technology and Operations Management, Erasmus University Rotterdam, Rotterdam, The Netherlands. [80]Faculty of Psychology, Universidad Rey Juan Carlos, Madrid, Spain. [81]Faculty of Psychology, Universidad Autónoma de Madrid, Madrid, Spain. [82]University of Göttingen, Göttingen, Germany. [83]Department of Psychology, University of Chicago, Chicago, IL, USA. [84]Erasmus School of Philosophy, Erasmus University Rotterdam, Rotterdam, The Netherlands. [85]Department of Political Science, Duke University, Durham, NC, USA. [86]Psychology of Language Department, University of Göttingen, Göttingen, Germany. [87]School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK. [88]School of Psychology, Queen's University Belfast, Belfast, UK. [89]School of Psychology, University of Kent, Canterbury, UK. [90]Leibniz Institute for Psychology (ZPID), Trier, Germany. [91]Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester, UK. [92]Department of Psychology, University of Cologne, Cologne, Germany. [93]School of Psychology, University of Sussex, Brighton, UK. [94]Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA. [95]Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands. [96]Independent researcher, Belfast, UK. [97]Jubail English Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Jubail, Saudi Arabia. [98]These authors contributed equally: Sam Parsons, Flávio Azevedo, Mahmoud M. Elsherif.
✉e-mail: flavio.azevedo@uni-jena.de

Published online: 21 February 2022
https://doi.org/10.1038/s41562-021-01269-4

References

1. Wilson, B. & Fenner, M. *Educ. Rev.* **47**, 1–4 (2012).
2. Gelman, A. & Loken, E. The garden of forking paths: why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. *stat.columbia.edu*, https://go.nature.com/3rJEoMH (2013).
3. Byrne, J. A. & Christopher, J. *FEBS Lett* **594**, 583–589 (2020).
4. Tierney, W. et al. *Organ. Behav. Hum. Decis. Process.* **161**, 291–309 (2020).
5. Kathawalla, U.-K., Silverstein, P. & Syed, M. *Collabra. Psychol.* **7**, 18684 (2021).
6. Crüwell, S. et al. *Z. Psychol.* **227**, 237–248 (2019).
7. Azevedo, F. et al. Introducing a Framework for Open and Reproducible Research Training (FORRT). Preprint at https://doi.org/10.31219/osf.io/bnh7p (2019).
8. Nosek, B. A. & Bar-Anan, Y. *Psychol. Inq.* **23**, 217–243 (2012).

Acknowledgements

We thank E. Fried and E. Uhlmann for helping to steer the discussion and reviewing some glossary terms.

Author contributions

S.P., F.A. and M.M.E. share co-first authorship. The order of the author list, with the exception of the first four authors and final author (that is, S.P., F.A., M.M.E., S.G. and B. Aczel), was determined via a random-number generator using R. S.P., F.A. and B. Aczel conceptualized the work. F.A. and S.G. curated data, undertook formal analysis and developed software. S.P. and F.A. performed investigations, developed methodology and administered the project. S.P., F.A., M.M.E., and S.G. performed validations. F.A. and S.G. generated visualizations. S.P., F.A., M.M.E., S.G., M.L.V., M.R.V., J.F.M., B. Arendt, J.T., M. Pownall, T.R.E., B. Szaszi, E.C., T.B.L., A.B.M., W.X.N., T.K., Y.Y., A.O'M., B.J.B., S.F., H.E.K., L.S., E.N., M.J., M. Paul, G.H.G., A.J.P., O.R., O.N.S., C.R.M., E.M.R., S.K.Y., A.B., E. A. Woodward, B.J.G., A.A.F., Q.X., A.G.C., C.L., C.T.K., F.D., R.M., B. Schmidt, B.G.F., M.Z., E.L.H., T.D., B.M.J.K., N.V., K.K., M.A.B., A.A., D.G.R., J.M., R.M.R., A.H., M.W., T.W., C.R.P., O.R.v.d.A., S.L.K.S., L.A.A., J.S.P., C.M.F.O., B.B.B., M.K.T., D.M., R.A.M., G.H.V., H.H., J.P.C., N.B., B.V., E.G.-G., Z.M.F., O.L., N.A.-A., Y.-F.Y., M.I., M.L., J.E.B., G.R., A.J.S.,

A.L., C.P., S.V., C.J.G., A.H.A.-H. and B. Aczel wrote the original draft. S.P., F.A., M.M.E., S.G., L.M., M.L.V., M.R.V., J.F.M., B. Arendt, M. Pownall, T.R.E., B. Szaszi, T.B.L., W.X.N., T.K., Y.Y., A.O'M., B.J.B., S.F., H.E.K., L.S., E.N., M.J., M. Paul, G.H.G., A.J.P., O.R., O.N.S., C.R.M., E.M.R., S.K.Y., A.B., A.R.B., B.J.G., A.A.F., Q.X., A.G.C., C.T.K., F.D., B. Schmidt, B.G.F., M.Z., E.L.H., T.D., B.M.J.K., G.F., K.K., M.A.B., A.A., D.G.R., J.M., R.M.R., A.H., M.W., T.W., C.R.P., O.R.v.d.A., S.L.K.S., L.A.A., J.S.P., C.M.F.O., B.B.B., M.K.T., D.M., R.A.M., H.H., J.P.C., N.B., E. A. Williams, T.B.R., S.R., B.V., B.J.I., M.K., Z.K., E.G.-G., Z.M.F., O.L., N.A.-A., K.I., Y.-F.Y., M.G.-H., M.I., K.J.F., M.L., V.A., J.E.B., A.A.A., S.A.-J., W.L., G.R., S.L.M., A.J.S., E.G.-P., A.L., C.P., A.T., S.V., M.C.L., C.J.G., J.K.W., A.H.A.-H., D.A.K. and B. Aczel reviewed and edited the manuscript.

Competing interests

The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-021-01269-4.