

10

11

Title: A Community-Sourced Glossary of Open Scholarship Terms

12 Authors:

13 Sam Parsons^{†1}, Flávio Azevedo^{‡2}, Mahmoud M. Elsherif^{‡3,4}, Samuel Guay⁵, Owen N. Shahim⁶,
14 Gisela H. Govaart^{7,8}, Emma Norris⁹, Aoife O'Mahony¹⁰, Adam J. Parker^{1,11}, Ana Todorovic¹,
15 Charlotte R. Pennington¹², Elias Garcia-Pelegrin¹³, Aleksandra Lazić¹⁴, Olly M. Robertson^{15,16},
16 Sara L. Middleton^{17,18}, Beatrice Valentini¹⁹, Joanne McCuaig^{20,21}, Bradley J. Baker^{22,23}, Elizabeth
17 Collins²⁴, Adrien A. Fillon²⁵, Tina B. Lonsdorf²⁶, Michele C. Lim¹, Norbert Vanek^{27,28}, Marton
18 Kovacs^{29,30}, Timo B. Roettger³¹, Sonia Rishi³², Jacob F. Miranda³³, Matt Jaquiere¹, Suzanne L.
19 K. Stewart³⁴, Valeria Agostini³², Andrew J. Stewart³⁵, Kamil Izydorzak³⁶, Sarah Ashcroft-
20 Jones¹, Helena Hartmann^{37,38}, Madeleine Ingham³², Yuki Yamada³⁹, Martin R. Vasilev⁴⁰, Filip
21 Dechterenko⁴¹, Nihan Albayrak-Aydemir⁴², Yu-Fang Yang⁴³, LaPlume A. Annalise⁴⁴, Julia K.
22 Wolska⁴⁵, Emma L. Henderson⁴⁶, Mirela Zaneva¹, Benjamin G. Farrar¹³, Ross Mounce⁴⁷,
23 Tamara Kalandadze⁴⁸, Wanyin Li³², Qinyu Xiao⁴⁹, Robert M. Ross⁵⁰, Siu Kit Yeung⁴⁹, Meng
24 Liu⁵¹, Micah L. Vandegrift⁵², Zoltan Kekecs^{29,53}, Marta K. Topor⁵⁴, Myriam A. Baum⁵⁵, Emily
25 A. Williams⁵⁶, Asma A. Assanee^{57,21}, Amélie Bret⁵⁸, Aidan G. Cashin^{59,60}, Nick Ballou⁶¹,
26 Tsvetomira Dumbalska¹, Bettina M. J. Kern⁶², Claire R. Melia^{63,64}, Beatrix Arendt⁶⁵, Gerald H.
27 Vineyard⁶⁶, Jade S. Pickering^{67,68}, Thomas R. Evans⁶⁹, Catherine Laverty³², Eliza A.
28 Woodward⁷⁰, David Moreau⁷¹, Dominique G. Roche⁷², Eike M. Rinke^{73,74}, Graham Reid¹,
29 Eduardo Garcia-Garzon⁷⁵, Steven Verheyen⁷⁶, Halil E. Kocalar⁷⁷, Ashley R. Blake²¹, Jamie P.
30 Cockcroft⁷⁸, Leticia Micheli⁷⁹, Brice Beffara Bret⁵⁸, Zoe M. Flack⁸⁰, Barnabas Szaszi⁸¹, Markus
31 Weinmann^{82,83}, Oscar Lecuona^{84,85}, Birgit Schmidt⁸⁶, William X. Ngiam⁸⁷, Ana Barbosa
32 Mendes⁸⁸, Shannon Francis³, Brett J. Gall⁸⁹, Mariella Paul⁹⁰, Connor T. Keating³², Magdalena
33 Grose-Hodge²¹, James E. Bartlett⁹¹, Bethan J. Iley^{92,93}, Lisa Spitzer⁹⁴, Madeleine Pownall⁵⁶,
34 Christopher J. Graham^{95,96}, Tobias Wingen⁹⁷, Jenny Terry⁹⁸, Catia M. F. d. Oliveira⁷⁸, Ryan A.
35 Millager⁹⁹, Kerry Fox⁸⁰, Alaa AlDoh⁹⁸, Alexander Hart⁵⁵, Olmo R. van den Akker¹⁰⁰, Gilad
36 Feldman⁴⁹, Dominik A. Kiersz¹⁰¹, Christina Pomareda³², Kai Krautter⁵⁵, Ali H. Al-Hoorie¹⁰²,
37 Balazs Aczel²⁹

38 Affiliations:

39 ¹Department of Experimental Psychology, University of Oxford, ²Department of
40 Communication, Friedrich Schiller University, ³Department of Psychology, University of
41 Birmingham, ⁴Blavatnik School of Government, University of Oxford, ⁵Department of
42 Psychology, University of Montreal, ⁶East Midlands School of Anaesthesia, ⁷Department of
43 Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig,
44 ⁸Charité – Universitätsmedizin Berlin, Einstein Center for Neurosciences Berlin, ⁹Department of
45 Health Sciences, Brunel University, ¹⁰School of Psychology, Cardiff University, ¹¹Department of
46 Experimental Psychology, Division of Psychology and Language Sciences, University College

47 London, ¹²School of Psychology, Aston University, ¹³Department of Psychology, University of
 48 Cambridge, ¹⁴Department of Psychology, University of Belgrade, ¹⁵Departments of Psychiatry
 49 and Experimental Psychology, University of Oxford, ¹⁶School of Psychology, Keele University,
 50 ¹⁷Department of Zoology, University of Oxford, ¹⁸Plant Sciences Department, University of
 51 Oxford, ¹⁹Faculty of Psychology and Educational Sciences, University of Geneva, ²⁰Department
 52 of Foreign Languages, Hongik University, ²¹Department of English Language and Applied
 53 Linguistics, University of Birmingham, ²²Department of Sport and Recreation Management,
 54 Temple University, ²³Mark H. McCormack Department of Sport Management, University of
 55 Massachusetts, ²⁴Department of Psychology, University of Stirling, ²⁵Department of Psychology,
 56 University of Aix-Marseille, ²⁶Institute for Systems Neuroscience, University Medical Center
 57 Hamburg-Eppendorf, ²⁷School of Cultures, Languages and Linguistics, University of Auckland,
 58 ²⁸Experimental Research on Central European Languages Lab, Charles University Prague,
 59 ²⁹Institute of Psychology, ELTE, Eotvos Lorand University, Budapest, Hungary, ³⁰Doctoral
 60 School of Psychology, ELTE Eotvos Lorand University, ³¹Department of Linguistics and
 61 Scandinavian Studies, Universitetet i Oslo, ³²School of Psychology, University of Birmingham,
 62 ³³Department of Psychology, University of Alabama, ³⁴School of Psychology, University of
 63 Chester, ³⁵Division of Neuroscience and Experimental Psychology, University of Manchester,
 64 ³⁶Faculty of Psychology, SWPS University of Social Sciences and Humanities, ³⁷Department for
 65 Cognition, Emotion, and Methods in Psychology, University of Vienna, ³⁸Netherlands Institute
 66 for Neuroscience, ³⁹Faculty of Arts and Science, Kyushu University, ⁴⁰Department of
 67 Psychology, Bournemouth University, ⁴¹Department of Cognitive Psychology, Institute of
 68 Psychology, Czech Academy of Sciences, ⁴²London School of Economics and Political Science,
 69 ⁴³Department of Psychology, University of Würzburg, ⁴⁴Rotman Research Institute, Baycrest
 70 Hospital (fully affiliated with the University of Toronto), ⁴⁵Department of Psychology,
 71 Manchester Metropolitan University, ⁴⁶School of Psychology, University of Surrey, ⁴⁷Arcadia
 72 Fund, ⁴⁸Department of Education, ICT and Learning, Østfold University College, ⁴⁹Department
 73 of Psychology, University of Hong Kong, ⁵⁰Department of Psychology, Macquarie University,
 74 ⁵¹Faculty of Education, University of Cambridge, ⁵²Open Knowledge Librarian, NC State
 75 University, ⁵³Department of Psychology, Lund University, ⁵⁴Department of Psychology,
 76 University of Surrey, ⁵⁵Research Group Applied Statistical Modeling, Saarland University,
 77 ⁵⁶School of Psychology, University of Leeds, ⁵⁷Taibah University, ⁵⁸Department of Psychology,
 78 University of Nantes, ⁵⁹Centre for Pain IMPACT, Neuroscience Research, ⁶⁰Prince of Wales
 79 Clinical School, University of New South Wales, ⁶¹Game AI Group, Queen Mary University of
 80 London, ⁶²Department of Cognition, Emotion, and Methods in Psychology, University of
 81 Vienna, ⁶³Department of Mental health and Social Work, Middlesex University, ⁶⁴Department of
 82 Psychology, Keele University, ⁶⁵Center for Open Science, ⁶⁶Department of Psychology,
 83 University of East London, ⁶⁷School of Psychology, University of Southampton, ⁶⁸Division of
 84 Neuroscience & Experimental Psychology, University of Manchester, ⁶⁹School of Human
 85 Sciences, University of Greenwich, ⁷⁰Department of Psychology, Oxford Brooke University,
 86 ⁷¹School of Psychology, University of Auckland, ⁷²Department of Biology, Carleton University,

87 ⁷³School of Politics and International Studies, University of Leeds, ⁷⁴Mannheim Centre for
88 European Social Research (MZES), University of Mannheim, ⁷⁵Faculty of Health, Universidad
89 Camilo José Cela, ⁷⁶Department of Psychology, Education and Child Studies; Erasmus
90 University Rotterdam, ⁷⁷Department of Educational Science, Muğla Sıtkı Koçman University,
91 ⁷⁸Department of Psychology, University of York, ⁷⁹Institute of Psychology, Leibniz University
92 Hannover, ⁸⁰School of Humanities and Social Science, University of Brighton, ⁸¹Institute of
93 Psychology, ELTE, Eotvos Lorand University, ⁸²Faculty of Management, Economics and Social
94 Sciences, University of Cologne, ⁸³Department of Technology and Operations Management,
95 Erasmus University Rotterdam, ⁸⁴Faculty of Psychology, Universidad Rey Juan Carlos, Madrid,
96 Spain, ⁸⁵Faculty of Psychology, Universidad Autónoma de Madrid, ⁸⁶University of Göttingen,
97 ⁸⁷Department of Psychology, University of Chicago, ⁸⁸Erasmus School of Philosophy. Erasmus
98 University Rotterdam, ⁸⁹Department of Political Sciences, Duke University, ⁹⁰Psychology of
99 Language Department, University of Göttingen, ⁹¹School of Psychology and Social Science,
100 Arden University, ⁹²School of Psychology, Queen's University Belfast, ⁹³School of Psychology,
101 University of Kent, ⁹⁴Leibniz Institute for Psychology (ZPID), ⁹⁵Division of Population Health,
102 Health Services Research and Primary Care, University of Manchester, ⁹⁶Scottish Health Action
103 on Alcohol Problems (SHAAP), ⁹⁷Department of Psychology, University of Cologne, ⁹⁸School
104 of Psychology, University of Sussex, ⁹⁹Department of Hearing and Speech Sciences, Vanderbilt
105 University, ¹⁰⁰Department of Methodology, Tilburg University, ¹⁰¹Independent, ¹⁰²Jubail English
106 Language and Preparatory Year Institute, Royal Commission for Jubail and Yanbu, Saudi Arabia

107

108 Correspondence: Flávio Azevedo (flavio.azevedo@uni-jena.de). ‡Shared co-first authorship.

109

110

111

112 **Standfirst:**

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

117

118 **Barriers to Open Scholarship Terminology**

119 Open Scholarship is an umbrella term referring to the endeavour to improve openness,
 120 integrity, social justice, diversity, equity, inclusivity and accessibility in all areas of scholarly
 121 activities. Open Scholarship extends the more widely used terms "Open Science" and "Open
 122 Research" to include academic fields beyond the sciences and academic activities.

123 Over the last decade, Open Scholarship has radically changed the way we think and
 124 discuss research and higher education. New concepts, tools, and practices have been developed
 125 and promoted, introducing novel terms or repurposing existing ones. These changes have
 126 increased the breadth but also the ambiguity of terminology, creating barriers to effective
 127 understanding and communication for novices and experts. Presently, certain terms such as
 128 *replicability* or *reproducibility* are well known but frequently used interchangeably or
 129 differentially among fields and disciplines; other terms are less known beyond a small circle of
 130 researchers, such as *CARKing*, *PARKing*, or *paradata*. Terms that become conventional within a
 131 given field often reflect the preferences of those with the platforms and privileges to determine
 132 academic discourse and, consequently, can act as a barrier to participation by those without such
 133 platforms. A similar barrier is that much academic language is contained within a ‘hidden
 134 curricula’, meaning these terms and practices are often used under the misplaced assumption that
 135 students, or those unfamiliar with an area, understand them.

136

137 **A Diversity of New Terminologies**

138 Terms associated with Open Scholarship are diverse in many aspects. Some are
 139 neologisms (i.e., newly coined) while others are reclamations of older terms (e.g. *p-hacking* and
 140 adversarial collaboration). Furthermore, frequent use of acronyms can hinder immediate
 141 understanding. For example, ORCID iD¹ refers to a persistent unique identifier for individuals in
 142 their role as creator or contributor. It enables linking digital documents and other contributions to
 143 their digital records, attributes credit, and resolves name ambiguities. Other terms use metaphors,
 144 such as the *Garden of Forking Paths*², which highlights the many alternative paths researchers
 145 can embark on when analysing data. These examples highlight the complexity of Open
 146 Scholarship terminology, which often assumes prior knowledge, making it difficult for
 147 individuals not versed in these terms to engage in ongoing conversations about these topics, thus
 148 excluding them from joining the discourse.

149

150 Communication across disciplines, and across dramatically varying levels of subject and
 151 technical expertise, can be extremely difficult. Challenges can arise in understanding scientific

152 texts when words with a historical meaning gain a new one in certain contexts. The term *paper*
153 *mills*³, for example, typically refers to factories devoted to manufacturing papers but also, in the
154 context of open scholarship, it denotes unethical for-profit organisations that create and publish
155 on-demand fraudulent scientific papers based on techniques such as fabrication of data and
156 plagiarism.

157 A similar challenge arises when the meaning and usage of certain terms differs between
158 (sub)disciplines and methodologies. For example, in social science fields the term
159 *preregistration* refers to an uneditable, timestamped version of a research protocol, whereas in
160 healthcare fields it refers to an accelerated course that qualifies students to fast-track into a
161 medical profession. As another example, social scientists understand the term *external validity* to
162 mean that the findings can be generalised to other contexts (different measures, people, places,
163 and times), while psychometricians regard it as the relationship of a psychological concept with
164 theoretically relevant extrinsic variables. *Creative destruction*⁴ is yet another example: in
165 economics, it refers to revolutionising the economic structure from within by destroying the old
166 system and replacing it with a new one. In psychology, a creative destruction approach to
167 replication means that replications can—in particular circumstances—be leveraged to not only
168 support or question the original findings, but also to replace weaker theories with stronger ones
169 that have greater explanatory power (by preregistering different theoretical predictions and
170 adding new measures, conditions and populations that facilitate competitive theory testing). As
171 interdisciplinary collaboration is growing and often required by many funding agencies and
172 stakeholders, this creates a potential for miscommunication and confusion when using such
173 terminology.

174 The clarification of scholarly terminology is a challenging endeavour. It should be built
175 on the insights of a community of experts with different perspectives and requires consensus
176 among the members across disciplines. As the breadth of these initiatives can be overwhelming,
177 digestible introductions to the language of Open Scholarship are needed⁵⁻⁷. In order to reduce
178 barriers to entry and understanding of Open Scholarship terminology, as well as to foster the
179 accessibility, inclusivity, and clarity of its language, a community-driven glossary using a
180 consensus-based methodology could help clarify terminologies and aid in the mentoring and
181 teaching of these concepts.

182

183 **The Open Scholarship Glossary Project**

184 The present glossary was developed by the FORRT⁷ community; an educational initiative
185 aiming to integrate open and reproducible research principles into higher education as well as
186 supporting educators and mentors to address related pedagogical challenges. The work has been
187 completed in three rounds. First, the lead team created an initial list of Open Scholarship related
188 terms and a structure for each term. Each term was required to have (1) a concise *definition*; with
189 supporting (2) *references*; (3) *related terms*; and (4) any applicable *alternative definitions*. The
190 present glossary has been developed using a crowdsourced methodology with the involvement of
191 over 100 contributors at various career stages and from a diverse range of disciplines, including

192 psychology, economics, neuroscience, information science, social science, biology, ecology,
 193 public health, and linguistics. In the second round, and in a dynamic and iterative crowdsourced
 194 process, members of the research community were invited through social media platforms or via
 195 organisations such as *ReproducibiliTea* to participate in the project. The community contributors
 196 suggested new terms, to which they provided main and alternative definitions, as well as
 197 reviewed and edited other terms iteratively throughout the project. We recorded these
 198 contributions, and this is reflected in our CRediT statement; all contributors were invited as
 199 authors on this manuscript. We considered definitions as ready for dissemination when they had
 200 been reviewed by a sufficient number of contributors (typically five or more) and reached
 201 consensus. Through this process, the community-driven glossary development procedure
 202 deliberately centred the Open Scholarship ethos of accessibility, diversity, equity, and inclusion.
 203

204 The project resulted in the drafting of more than 250 terms. In Table 1, we present an
 205 abbreviated set of 30 terms that represent the plurality of terms for the broader Open Scholarship
 206 concepts (see <https://forrt.org/glossary> for the complete glossary, including key references and
 207 links to related terms, as well as a more detailed explanation of the project's mission and goals).
 208 The FORRT Glossary project is licensed under a CC BY NC SA 4.0 license. The present
 209 glossary is the 1.0 version. The version-controlled source code of the new releases of the
 210 complete Glossary is archived on FORRT's [website](#), [GitHub](#), [OSF](#), and [Zenodo](#), wherein new
 211 releases will also be stored. We set up a system allowing for continuous improvement, extension,
 212 and updating from community feedback and involvement. Versioning will also allow the study
 213 of the evolution of the terminologies.
 214

215 **Table 1 Open Scholarship Glossary 1.0 Examples (abbreviated version)**

216 *Note:* The complete glossary, including key references and links to related terms, can be found in
 217 Supplementary Information and is available at <https://forrt.org/glossary>. Minor modifications
 218 were made to comply with editorial requirements.

Analytic Flexibility	Analytic flexibility is a type of researcher degrees of freedom that 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	A tongue-in-cheek expression intended to raise awareness of the lack of diverse voices in open science, in addition to the presence of behavior 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 criticism of a design as one that you would have made in advance of the results being known. It usually forms a reaction or criticism to 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 data set. 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 (i.e., using different operationalisations, 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 generalised (e.g., only within certain contexts, with certain samples, using certain measurement approaches) towards evaluating and advancing theory.
Creative destruction approach	Replication efforts should seek not just 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 (e.g., 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	The Contributor Roles Taxonomy (CRediT; https://casrai.org/credit/) is a high-level taxonomy used to indicate the roles typically adopted by contributors to scientific scholarly output. There are currently 14 roles that describe each contributor’s specific contribution to the scholarly output.

	<p>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, Writing – review & editing.</p>
Decolonisation	<p>Coloniality can be described as the naturalisation 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 to understand their persistence and to reconstruct the norms and values of a given domain. In an academic setting, decolonisation refers to the rethinking of the lens through which we teach, research, and co-exist, so that the lens generalises beyond Western-centred and colonial perspectives. Decolonising 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 whilst holding a perspective that is separate from the socially dominant perspective. Also, by including, not rejecting, an individuals' internalised norms and taboos from the specific colony.</p>
Direct replication	<p>As 'direct replication' does not have a widely-agreed technical meaning nor 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 generalised effects (phenomena). How closely a replication replicates an original study is often a matter for debate, often with</p>

	<p>differences being cited as hidden moderators of effects. Furthermore, there can be debate over the relevant importance of technical equivalence (i.e., using identical materials) versus psychological equivalence (i.e., realizing the identical psychological conditions) to the original study. For example, consider a study on Trust in the US- President conducted in 2018. A technical equivalent replication would use Trump as stimulus (he was president in 2018) a psychological equivalent study would use Biden (he is the current president).</p>
External Validity	<p>Whether the findings of a scientific study can be generalized to other contexts outside 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 (e.g., an artificial laboratory setting or a non-representative sample). Alternative definition: In Psychometrics, the degree of evidence that confirms the relations of a tested psychological construct with external variables.</p>
FAIR principles	<p>Describes making scholarly materials Findable, Accessible, Interoperable and Reusable (FAIR). ‘Findable’ and ‘Accessible’ are concerned with where materials are stored (e.g. in data repositories), while ‘Interoperable’ and ‘Reusable’ focus on the importance of data formats and how such formats might change in the future.</p>
Garden of forking paths	<p>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’ (Gelman and Loken, 2013, p. 6). 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” (Gelman and Loken, 2013, p. 1). 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 (e.g. p-hacking and fishing expeditions) versus non-intentional research practices that can, potentially, have the same effect despite not having intent to corrupt their results. The</p>

	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 which could have been taken (had other decisions been made).
HARKING	A questionable research practice termed ‘Hypothesizing After the Results are Known’ (HARKing). HARKing has been defined as a post hoc hypothesis which is either 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	The set of evaluation and reward mechanisms (explicit and implicit) for scientists and their work. Incentivised 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 telos of science, and therefore do not promote rigorous scientific output. Initiatives like DORA aim to reduce the field’s dependency on evaluation criteria such as journal impact factors in favor 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) where diversity simply indicates a wide range of backgrounds, perspectives, and experiences, efforts to increase inclusion go further to promote engagement and equal valuation among diverse individuals, who might otherwise be marginalized. Increasing inclusivity often involves minimising the impact of, or even removing, systemic barriers to accessibility and engagement.
Metadata	Structured data that describes and synthesises other data. Metadata can help find, organize, and understand data. Examples of metadata include creator, title, contributors, keywords, 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, the sample and data collection process. <i>Alternative definition</i> : 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 impact 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 includes all scholarly activities that are not solely limited to research such as teaching and pedagogy.
Open Science	An umbrella term reflecting 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 behaviors 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	An organisation 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	Exploiting techniques that may artificially increase the likelihood of obtaining a statistically significant result by meeting the standard statistical significance criterion (typically $\alpha = .05$). For example, performing multiple analyses and reporting only those at $p < .05$, selectively removing data until $p < .05$, selecting variables for use in analyses based on whether those parameters are statistically significant.
Papermill	An organization that is engaged in scientific misconduct wherein multiple papers are produced by falsifying or fabricating data, e.g. by editing figures or numerical data or plagiarizing written text. A papermill 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 (e.g. through plagiarism software).
Paradata	Data that are captured about the characteristics and context of primary data collected from an individual - 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, 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 the practice where researchers complete an experiment (possibly with infinite re-experimentation) before preregistering. This practice invalidates the purpose of preregistration, and is one of the QRPs (or, even scientific misconduct) that try to gain only "credibility that it has been preregistered.
Preregistration	The practice of publishing the plan for a study, including research questions/hypotheses, research design, data analysis before the data has 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 (e.g., a repository) so that it can be publicly shared with others (possibly after an embargo period). Preregistration provides a 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	An umbrella term, used differently across fields, covering concepts of: direct and conceptual replication, computational reproducibility/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 artifacts; 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	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 regenerate 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	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 (IPA) 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 conducted 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.

221 **Summary**

222 This glossary is a first step in creating a common language for these concepts, facilitating
223 discussions about the strengths and weaknesses of different Open Scholarship practices, and
224 ultimately helping to build a stronger research community⁸.

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

229

230 **Competing interests**

231 The other authors declare no competing interests.

232

233 **Data and materials availability**

234 All materials produced are available in its entirety at <https://forrt.org/glossary>

235

236 **Acknowledgements**

237 We would like to thank Dr. Eiko Fried and Professor Dr. Eric Uhlmann for helping steer
238 the discussion and reviewing some Glossary terms.

239

240 **Author contributions**

241 **Conceptualization:** S.P., F.A., and B. Aczel.

242 **Data curation:** F.A. and S.G.

243 **Formal analysis:** F.A. and S.G.

244 **Investigation:** S.P. and F.A.

245 **Methodology:** S.P. and F.A.

246 **Project administration:** S.P. and F.A.

247 **Software:** F.A. and S.G.

248 **Validation:** S.P., F.A., M.M.E., and S.G.

249 **Visualization:** F.A. and S.G.

250 **Writing - original draft:** S.P., F.A., M.M.E., S.G., M.L.V., M.R.V., J.F.M., B. Arendt, J.T., M.
 251 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.,
 252 H.E.K., L.S., E.N., M.J., M. Paul, G.H.G., A.J.P., O.M.R., O.N.S., C.R.M., E.M.R., S.K.Y.,
 253 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,
 254 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.,
 255 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.d.O., B.B.B., M.K.T., D.M.,
 256 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.,
 257 J.E.B., G.R., A.J.S., A.L., C.P., S.V., C.J.G., A.H.A.-H., and B. Aczel

258 **Writing - review & editing:** S.P., F.A., M.M.E., S.G., L.M., M.L.V., M.R.V., J.F.M., B.
 259 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.,
 260 H.E.K., L.S., E.N., M.J., M. Paul, G.H.G., A.J.P., O.M.R., O.N.S., C.R.M., E.M.R., S.K.Y.,
 261 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.,
 262 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.,
 263 O.R.v.d.A., S.L.K.S., L.A.A., J.S.P., C.M.F.d.O., B.B.B., M.K.T., D.M., R.A.M., H.H., J.P.C.,
 264 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.,
 265 Y.-F.Y., M.G.-H., M.I., K.F., M.L., V.A., J.E.B., A.A.A., S.A.-J., W.L., G.R., S.L.M., A.J.S.,
 266 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

267 **List of Supplementary Materials:**

268 The unabbreviated version of the Open Scholarship Glossary can be found at the supplementary
 269 materials.

270

271 **References:**

- 272 1. Wilson, B. & Fenner, M. Open researcher & contributor ID (ORCID): solving the name
 273 ambiguity problem. *Educ. Rev* **47**, 1–4 (2012).
- 274 2. Gelman, A. & Loken, E. The garden of forking paths: Why multiple comparisons can be a
 275 problem, even when there is no “fishing expedition” or “p-hacking” and the research
 276 hypothesis was posited ahead of time. *Dep. Stat. Columbia Univ.* **348**, (2013).
- 277 3. Byrne, J. A. & Christopher, J. Digital magic, or the dark arts of the 21st century—how can
 278 journals and peer reviewers detect manuscripts and publications from paper mills? *FEBS Lett.*
 279 **594**, 583–589 (2020).
- 280 4. Tierney, W. *et al.* Creative destruction in science. *Organ. Behav. Hum. Decis. Process.* **161**,

- 281 291–309 (2020).
- 282 5. Kathawalla, U.-K., Silverstein, P. & Syed, M. Easing Into Open Science: A Guide for
283 Graduate Students and Their Advisors. *Collabra Psychol.* **7**, (2021).
- 284 6. Crüwell, S. *et al.* Seven Easy Steps to Open Science. *Z. Für Psychol.* **227**, 237–248 (2019).
- 285 7. Azevedo, F., *et al.* Introducing a Framework for Open and Reproducible Research Training
286 (FORRT). (2019) doi:10.31219/osf.io/bnh7p.
- 287 8. Nosek, B. A. & Bar-Anan, Y. Scientific utopia: I. Opening scientific communication. *Psychol.*
288 *Inq.* **23**, 217–243 (2012).
- 289
- 290