Open science for better research

Sara Garofalo
University of Bologna

What is Open Science?

Open science is an umbrella term encompassing fundamental ideas and practices concerning how science should be conducted and implemented. Core ideas include openness, transparency, rigor, reproducibility, and replicability, which in turn involve practices like openly creating, accessing, and sharing all resources contained in the research process (Claesen et al., 2021; Kathawalla et al., 2021; Zandonella Callegher & Massidda, 2022). The foundational principle of good methodological practice under open science requires researchers to be as transparent and open as possible throughout all phases of the research cycle. This transparency is necessary so that readers can fully and appropriately evaluate the work presented (Kathawalla et al., 2021). This required level of transparency allows readers to fully and appropriately evaluate the work being presented. Open science advocates frequently motivate their practices by appealing to Merton's norms or imperatives (Merton, 1973; Wagenmakers et al., 2021). These four philosophical principles, summarized by the acronym CUDOS, are foundational to the optimal functioning of science. They dictate that scientists should assess claims based purely on the evidence (Universalism), ensure that the products of science are the shared property of the community (Communalism), remain free from self-interest or personal gain (Disinterestedness), and subject all claims to systematic scrutiny and verification (Organized Skepticism). The emergence of issues like the 'replication crisis' across various disciplines highlighted how modern scientific practices have drifted away from these foundational ideals. Open Science is configured as a systematic attempt to re-align scientific practice with these imperatives through the adoption of concrete practices. The primary goal of Open Science is therefore not just to be 'good science' a priori, but to provide the necessary tools and mechanisms to ensure that every step of the research cycle is verifiable. This operational transparency rests on four interconnected pillars: Open Access (open access to publications), Open Data (data sharing), Open Methods/Code (transparency of materials and code), and Open Evaluation (open review).

The reproducibility crisis and the Open Science revolution

Scientific research has faced increasing concerns regarding the credibility, statistical rigor, and reproducibility of many published empirical findings. This situation led to a widely discussed crisis of confidence, commonly referred to as the "reproducibility crisis" in many fields of science, inspiring a variety of methodological reforms aimed at increasing the quality of confirmatory empirical research. The crisis is typically understood as the insight that the extant published literature contains a worrying amount of unreplicable, false-positive findings. Scientific research relies on the distinction between hypothesis generation (postdiction), which involves using existing observations, and hypothesis testing (prediction), which involves using new observations (Nosek et al., 2018). When this distinction is not respected in practice, and postdiction is mistaken for prediction, the credibility of research findings is reduced (Nosek et al., 2018). This issue, coupled with inherent biases in human reasoning, such as hindsight bias, as well as scientific misconduct, has led to increasing concerns about the credibility, reproducibility, and statistical rigor of empirical findings (Munafò et al., 2017; Open Science Collaboration, 2015; Pennington & Heim, 2022). In psychology, the empirical evidence for this crisis stems from a large-scale replication effort. In 2015, the Open Science Collaboration conducted a collaborative study to obtain an initial estimate of the reproducibility of psychological science by performing direct replications of 100 published studies to check whether consistent results could be found. The study showed that, while 97% of the original studies reported statistically significant results, only 36% of the replications yielded statistically significant findings. Moreover, the mean effect size of the replication effects was only half the magnitude of the original effect sizes. A complementary analysis found a 47.4% replication success rate when testing whether the original effect size was contained within the replication's 95% confidence interval. Critics argued that the OSC's method may have severely misjudged the actual rate of replication because many replication studies differed methodologically from the originals (Gilbert et al., 2016). However, the OSC authors defended their protocols, noting that in some cases, the original authors had recommended or endorsed the alleged methodological differences (Open Science Collaboration, 2016). A major explanation for irreproducibility is the strong publication bias that favors the publication of positive or statistically significant results. Classically, studies that identify a relationship or treatment effect are more likely to be published than those reporting null results. This practice has roots in the statistical models most frequently used for hypothesis testing (e.g., null hypothesis significance testing, NHST). However, such an incentive structure ultimately rewarded research quantity over quality, thereby motivating researchers to focus exclusively on obtaining and reporting statistically significant results, which could potentially impact the accuracy of the findings. Replication studies have indeed confirmed that scientific misconduct and questionable research practices, such as HARKing (Hypothesizing After the Results are Known) and p-hacking (the practice of manipulating data, analyses, or statistical tests until they produce a statistically significant result) played an important role in the reproducibility crisis (Chambers & Tzavella, 2022; Munafò et al., 2017; Open Science Collaboration, 2015; Pennington & Heim, 2022). Fortunately, the influence of selective inference is detectable and addressable through transparency in the research process (Nosek et al., 2018). The response to this crisis was a strong desire within the scientific community to better understand the conceptual, methodological, and analytical choices made across the entire research cycle, in order to enhance knowledge and facilitate a more accurate assessment of credibility. The so-called "open science movement" thus emerged as a necessary response to these issues, aiming to promote transparency, accessibility, credibility, and reproducibility within scientific research (Kathawalla et al., 2021).

Open Science in practice

Open science practices refer to a set of principles and actions that ensure research is carried out and communicated with maximal transparency, accessibility, and rigor. For example, open science includes adopting writing standards aligned with established best practices (Appelbaum et al., 2018; Gernsbacher, 2018; Levitt et al., 2018). It also encompasses several practical components that shape everyday research workflows, such as responsible data sharing; the use of preregistrations, registered reports, and preprints; the sharing of materials and code; open access to publications; and, increasingly, forms of open peer review that make editorial processes more transparent. The following paragraphs will explore some of the most relevant practices for daily research.

Data sharing

The objective of data sharing is to make the de-identified dataset utilized for a project accessible to other researchers. This practice is generally rated as having medium difficulty, requiring foresight regarding consent during the design phase and effective organization during dissemination. It is insufficient to merely state that data are available upon request (Kathawalla et al., 2021). Transparent data sharing offers substantial benefits for cumulative science and for researchers themselves (Kathawalla et al., 2021), as it enables other scientists to verify the accuracy and quality of published work by reproducing original analyses and conducting additional robustness tests or fitting alternative models; it fosters knowledge accumulation by allowing datasets to be reused for new research questions or incorporated into meta-analyses; and it provides credit and professional advantages, since shared data increase article citation impact and can earn researchers formal recognition (Kathawalla et al., 2021).

FAIR data The Findable, Accessible, Interoperable, Reusable (FAIR) guiding principles represent a crucial framework designed to improve the transparency, efficiency, and impact of research by governing scientific data management and stewardship (Boeckhout et al., 2018; Sadeh et al., 2023). The principles emerged from a 2014 workshop held in Leiden, Netherlands, stemming from discussions among diverse stakeholders representing academia, industry, funding agencies, and publishers (Wilkinson et al., 2016). Subsequently, they were formalized and promoted by the Future of Research Communications and e-Scholarship (FORCE11) community (Wilkinson et al., 2016). Crucially, while there were earlier similar initiatives, the core distinction of FAIR is its specific emphasis on enhancing the ability of machines to autonomously find and use digital resources (referred to as machine-actionability) (Jacobsen et al., 2020; Wilkinson et al., 2016). This focus aims to maximize knowledge discovery and reuse by allowing computational agents to automatically discover, access, integrate, and analyze appropriate data at scale (Wilkinson et al., 2016). Although the principles deliberately avoid dictating specific technology choices, providing guidance rather than prescriptive standards, they are not themselves a standard or specification (Hasnain & Rebholz-Schuhmann, 2018; Jacobsen et al., 2020; Mons et al., 2017), and they allow individual stakeholder communities the freedom to define their own solutions (Jacobsen et al., 2020). Adherence to these principles is key to maximizing data value and realizing the scientific benefits of integration and reuse (Kush et al., 2020; Wilkinson et al., 2016). The FAIR acronym encompasses four foundational principles, which are further elaborated by 15 measurable guiding sub-principles (Jacobsen et al., 2020; Wilkinson et al., 2016). In practice, to achieve these goals, researchers must implement specific steps and utilize appropriate tools (Dunning et al., 2017; Jacobsen et al., 2020; Kush et al., 2020; Martone, 2024; Martone et al., 2018; Wilkinson et al., 2016):

- **Findability (F)** requires assigning a globally unique and persistent identifier (PID). For example, depositing a dataset into a recognized repository, such as the Open Science Framework (OSF), ensures the generation of a Digital Object Identifier (DOI) that persistently links to the data and its metadata.
- **Accessibility (A)** means the data are retrievable via a standardized process, even if authorization is required. Practically, this is achieved by using an open, universally implementable communication protocol, such as HTTP.
- **Interoperability (I)** addresses the requirement that metadata uses a formal, shared language to allow seamless integration across systems. For complex data, communities should adopt standards like Common Data Elements (CDEs) or

formal ontologies; for instance, BioPortal provides programmatic access to curated domain ontologies that aid in standardizing terminology.

- **Reusability (R)** necessitates richly described metadata and, critically, the inclusion of a clear data usage license to avoid legal impediments to reuse. Common examples are assigning a Creative Commons (CC) licence, detailing how the data was generated using structured frameworks such as PROV-Template, meet domain-relevant community standards, where they exist, such as the Minimum Information About a Microarray Experiment (MIAME) standard in genomics.

Of note, adopting FAIR data sharing practices is increasingly required or encouraged by major funders (Jacobsen et al., 2020; Kush et al., 2020; Wilkinson et al., 2016), with agencies like the NIH and the European Commission strengthening policies that embed FAIR principles into data management (Kush et al., 2020; Martone et al., 2018). Programs such as Horizon 2020 explicitly mandate FAIR-aligned data practices (Boeckhout et al., 2018; Dunning et al., 2017), and NIH now requires researchers to specify data-sharing venues in their plans (Martone, 2024). Consistent data archiving can also improve funding competitiveness (Sadeh et al., 2023), supporting better long-term data stewardship and reuse (Sadeh et al., 2023; Wilkinson et al., 2016).

Ethical and legal requirements Legal restrictions on data sharing must be considered, and researchers should always ensure they comply with their local legal requirements (Zandonella Callegher & Massidda, 2022):

- Informed consent and data permission: researchers must obtain both participants' consent to participate in the study and their explicit permission to share their research data. These are two different things.
- Best practice for permission: the best practice for seeking permission to share data is to do so after the completion of research activities.
- This ensures that participants are fully aware of the study's procedures and what they are being asked to share.
- Data protection: when sharing data, special attention must be paid to privacy rules.

Pre-registering hypotheses

Understanding the distinction between testing hypotheses and exploring data is central to transparent and credible research practice. Confirmatory analyses test predictions that were specified in advance, whereas exploratory analyses generate new hypotheses or identify unexpected patterns in the data (Claesen et al., 2021; Heers, 2020; Nosek et al., 2018). Both approaches are valuable, but they serve different purposes and must be communicated clearly to avoid misleading interpretations. This distinction is at the heart of many open science practices, which aim to prevent ambiguity, reduce questionable research practices (such as HARKing, Hypothesizing After the Results are Known), and strengthen the interpretability of findings. The following sections introduce two important procedural tools (Preregistration and Registered Reports) that support researchers in clearly delineating confirmatory and exploratory work. These practices help safeguard the integrity of hypothesis testing while still allowing space for productive exploration.

Preregistration (PR) Preregistration requires researchers to define their research questions and analysis plan before collecting data or before accessing an existing dataset (Henderson, 2022; Kathawalla et al., 2021; Nosek et al., 2018; Stewart et al., 2020). To do so, researchers are required to formally document a study's research plan on a public, time-stamped repository before data collection or observation (Claesen et al., 2021; Kathawalla et al., 2021; Stewart et al., 2020). Preregistration plans vary in specificity. Some are brief, describing only the research questions, hypotheses, and tests, whereas others include highly detailed elements such as analysis scripts based on simulated data (Kathawalla et al., 2021; Sarafoglou et al., 2022). In all cases, they require careful planning, as it should be written in sufficient detail to facilitate a fair comparison or adherence check (Claesen et al., 2021). Structured templates, such as those used in the OSF Preregistration format (https://osf.io/registries/), encourage researchers to provide the necessary level of detail and can constitute a practical and useful guide on how to write a pre-registration. These structured templates promote more comprehensive reporting on topics such as inference criteria and the management of missing data (Bakker, Veldkamp, van Assen, et al., 2020). The OSF offers various templates suited to different research contexts, including experimental designs and secondary data analysis (Kathawalla et al., 2021), and simple preregistration forms are also available through AsPredicted.org (Nosek et al., 2018). Preregistrations on OSF can be saved as drafts and locked once officially registered. Researchers may choose to make their preregistration immediately public or to place it under embargo for a specified period of up to four years (Kathawalla et al., 2021; Stewart et al., 2020). Importantly, preregistration is often described as a "plan, not a prison" (Henderson, 2022). Deviations from the preregistered protocol may be necessary due to honest mistakes, methodological limitations, or reviewer suggestions (Claesen et al., 2021; Henderson, 2022). However, to maintain transparency, all deviations must be disclosed, and exploratory analyses (i.e., those not initially planned or not part of the initial hypotheses) must be reported as such (Claesen et al., 2021; Stewart et al., 2020). Without disclosure, readers assume that the study was conducted exactly as planned (Claesen et al., 2021).

Registered Reports (RR) Registered Reports represent a publishing model that shifts the basis of publication decisions away from away from results and focuses on methodological rigor and conceptualization (Chambers, 2013; Chambers & Tzavella, 2022; Nosek, 2014). This format is generally considered more challenging than standard preregistration due to the additional time, coordination, and peer review involved; however, it offers significant advantages in both scientific accuracy and publication success. The Registered Report model encompasses a two-stage peer review process (Chambers & Tzavella, 2022; Frings, 2021; Henderson, 2022; Kathawalla et al., 2021; Stewart et al., 2020):

Stage 1 - Proposal Submission and Triage. Authors submit a detailed Stage 1 proposal encompassing the introduction, methods, and analysis plan prior to data collection. This proposal is peer-reviewed for the rigor and quality of the study design. Submissions must first pass an initial triage assessment. The submission must be comprehensive, including the hypotheses, sampling plan (such as a statistical power analysis or equivalent rationale), and planned contingencies. The review process focuses centrally on the precision of and adherence to the preregistered plan.

In-Principle Acceptance (IPA). A successful Stage 1 review results in an IPA, which is a guarantee of publication regardless of the study's results (including null findings), provided the authors adhere faithfully to the approved protocol.

Stage 2 - Adherence and Conclusion Check. After data collection, the full manuscript (including results and discussion) is submitted for Stage 2 review, which focuses strictly on confirming adherence to the accepted protocol and ensuring the conclusions are justified by the evidence. If reviewers find flaws in the protocol that were missed at Stage 1, the manuscript is generally not rejected on that basis, and editors cannot usually require authors to conduct extra studies.

Where and how to publish Preregistrations and Registered Reports **Preregistration** in its "standard" form (e.g., via OSF or similar platforms) is entirely voluntary and does not require any formal agreement from journals in advance. This means that any journal, regardless of whether it explicitly mentions preregistration in its author guidelines, can in principle accept a manuscript that reports having preregistered its hypotheses, methods, or analysis plan, as long as the authors provide a link or DOI to the preregistered document. In other words, authors are always free to preregister their study and later disclose this information in the submitted manuscript; the journal does not need to pre-approve or host the preregistration. A growing number of journals in psychology and related fields now encourage preregistration more directly, reflecting the increasing integration of open science practices into mainstream publishing. For instance, journals such as Psychological Science, Advances in Methods and Practices in Psychological Science (AMPPS), Nature Human Behaviour, PLOS ONE, and Collabra:

Psychology explicitly recommend preregistration in their author guidelines and/or recognize it through badges or dedicated sections for transparency statements. Beyond simple preregistration, an expanding set of journals also offer the **Registered Reports** format, in which peer review occurs before data collection, and the journal commits to publishing the study regardless of results if the preregistered protocol is followed. Psychology was an early adopter of this format, and many influential journals now support it. Examples include Cortex, Royal Society Open Science, BMC Psychology, European Journal of Personality, and Journal of Cognition. These journals provide structured submission pathways for Registered Reports and are signatories to the broader movement promoting transparency and reproducibility.

In addition to journal-based pathways, there are non commercial, researcher driven initiatives, such as Peer Community In (PCI), which offer another flexible and transparent route. Through its platform, PCI Registered Reports (PCI-RR), authors can submit a Stage 1 protocol for peer review independently of traditional publishing houses. Once PCI RR issues a positive Stage 2 recommendation, authors may publish their accepted manuscript in any "PCI friendly" journal (two examples being Experimental Psychology and Addiction Research & Theory) without further scientific peer review (Frings, 2021; Pennington & Heim, 2022). Because PCI is non-commercial and researcher-run, it provides an appealing option for authors seeking both methodological integrity and flexibility in publication venue.

All in all, the expanding availability of preregistration, Registered Reports, and community-driven peer review initiatives reflects a progressive transformation in the research ecosystem. These options give researchers more control over their scientific output, enhance transparency, and contribute to a more credible, reproducible, and equitable scientific record. Importantly, this shift signals that the field increasingly values strong study design, transparent planning, and methodological rigor over purely "positive" or novel results.

Open access publishing

While in earlier centuries (17th–19th), scientific societies circulated publications more informally or through membership, for most of the 20th and 21st centuries, scientific publishing has been dominated by subscription-based models, where access to research articles required payment either by readers, universities, or authors. This subscription-based system has shaped the circulation of knowledge, often limiting the reach and impact of scientific findings. Against this backdrop, Open Access (OA) publishing emerged as an alternative model aimed at removing financial and legal barriers to scholarly communication. Open Access refers to a set of principles and practices that ensure research outputs are freely available to anyone, a feature now recognized as foundational to open science alongside transparency, credibility, and reproducibility (Kathawalla et al., 2021). A core idea of good open science practice is to be as transparent and open as possible throughout the entire research cycle, enabling readers to fully and appropriately evaluate the work presented (Kathawalla et al., 2021). To support such openness, OA publications are often released under permissive licenses (most commonly the Creative Commons Attribution license - CC BY; https://creativecommons.org/), which allow others to copy, distribute, and reuse the material provided that the original author and source are credited (Bakker, Veldkamp, van den Akker, et al., 2020; Claesen et al., 2021; Ofosu & Posner, 2021).

In the contemporary publishing landscape, journals offer several models through which authors can make their work openly accessible.

The only fully open option is an increasingly recognized but less common alternative known as ***Diamond or Platinum Open Access***, where both reading and publishing are free. These journals rely on funding from universities, scholarly societies, or consortia rather than author fees. While they embody the ideals of open access, their long-term sustainability depends heavily on institutional support.

Gold Open Access arguably represents the second-best option. In gold journals, all published articles are immediately and freely available to anyone, typically under permissive Creative Commons licenses such as CC BY. Because these journals do not rely on subscription income, authors (or their funders or institutions) usually cover the costs through Article Processing Charges, which range from under a thousand to several thousand dollars.

Green Open Access, or self-archiving, offers a third route in which authors deposit a version of their manuscript (often the accepted manuscript rather than the final typeset version) in a repository. While this option is free and significantly expands access, it is limited by embargoes, version restrictions, and the fact that the publisher's final formatted article often remains behind a paywall. For this reason, green routes support openness but do not fully replace the need for gold open access.

By contrast, the so-called ***Hybrid Open Access*** occupies an ambiguous middle ground and is not considered true open access by many in the scholarly community. Hybrid journals continue to operate on a subscription model but offer authors the option to pay an APC to make their individual article openly accessible. The rest of the journal remains paywalled, and publishers retain subscription revenue while simultaneously collecting open-access fees. This "double dipping" has led to criticism and to the growing consensus that hybrid models do not align with the principles of open access, even if individual articles can be made openly available within them. Hybrid journals are nonetheless important to describe because they set the stage for the rise of **transformative agreements**. These institutional contracts, such as read-and-publish or publish-and-read deals, aim to shift subscription spending toward open access by bundling reading rights with the costs of publishing articles openly. Under such agreements, authors at participating institutions can publish open access in gold or hybrid journals without paying APCs themselves. The broader goal is to transition the financial model away from subscriptions and toward full open access, ideally reducing the reliance on hybrid mechanisms over time. Taken together, these models illustrate the transition from a historically paywalled system toward a more open and equitable one, with gold open access standing as the only fully open path, hybrid models providing a partial and imperfect bridge, and transformative agreements aiming to restructure the economics of publishing to support a more genuinely open future.

Preprints Other forms of open access publishing exist, including the rapid sharing of manuscripts via preprints, which can be posted before submission, while under review, or as an author formatted version of an accepted article (Kathawalla et al., 2021).

A preprint is an author-formatted manuscript that is posted publicly on an open-access repository (e.g., OSF Preprints). The term mostly refers to a version of a manuscript made publicly available prior to being submitted for peer-review, although the term is sometimes also used to refer to manuscripts under review or author-formatted versions of accepted articles (Kathawalla et al., 2021; Zandonella Callegher & Massidda, 2022). Preprints enable the rapid dissemination of findings, allowing researchers to receive wider feedback outside of the formal peer review process, which can help improve a paper by identifying major flaws prior to submission. Moreover, they help authors establish priority and protect against scooping concerns by creating a temporal record (date/time stamp) (Kathawalla et al., 2021; Moshontz, 2018; Moshontz et al., 2020). However, a critical caution regarding preprints is that, by definition, they may be posted prior to peer review. Therefore, unless the preprint explicitly indicates that it is a final version that has passed peer review, it should not be treated as finalized, verified scientific knowledge (Henderson, 2022; Kathawalla et al., 2021).

Useful and practical guides

Tools, manuals and papers Several sources provide explicit guidance and resources for implementing open science practices, including recommendations on which platforms, tools, or software to use for various needs.

Kathawalla, Silverstein, & Syed (2021) serve as a roadmap to help graduate students and their advisors engage in open science practices. This paper presents eight suggested open science practices, utilizing the format of 'what, why, how, and worries' for each topic. It includes specific ideas on how to approach conversations with advisors and suggests ways to integrate these practices into graduate school. The paper discusses practices across the research cycle (conceptualization, design, analysis, reporting, dissemination), including journal clubs, project workflow, preprints, reproducible code, data sharing, transparent writing, preregistration, and registered reports. It also offers specific resources on how to engage with each behavior and points to a companion OSF project page, which contains links to video tutorials and step-by-step guides.

Zandonella Callegher & Massidda (2022), can be used to learn how to: share materials using the Open Science Framework (OSF); organize project files and data in a structured and documented Repository; write readable code using a Functional Style approach; use Git and GitHub for tracking changes; manage the Analysis Workflow pipeline with dedicated tools; and create Dynamic Documents. This manual provides recommendations and guidelines that are useful for any programming language, although the examples are based on the R programming language. It recommends gradually building a reproducible workflow step-by-step, with each chapter focusing on an independent part of the process.

Moshontz, Binion, Walton, Syed, & Brown (2020) includes a detailed description of how and why researchers should post preprints.

Meyer (2018) is recommended for readers interested in the complexities associated with data sharing.

Soderberg (2018) provides instructions on how to create an OSF account, set up a project, add collaborators, upload files, and explore additional capabilities of an OSF project.

Kiyonaga & Scimeca (2019), referenced as "Practical Considerations for Navigating Registered Reports," is a helpful practical guide for authors pursuing the Registered Report format.

Henderson (2022) includes specific practical resources like a checklist of items to include when writing up the results of preregistered research and advice on when to apply for ethics. It also provides guidance on writing a Registered Report, including a template that guides authors through the key elements.

Gernsbacher (2018) provides excellent recommendations that are current and consistent with best practice in open science for transparent manuscript writing.

Kern & Gleditsch (2017) explore how pre-registration and pre-analysis plans can foster transparency in qualitative research and provide a pre-registration template for such purposes.

Open science initiatives **Reproducibility Networks** (RNs; https://reproducibility.global/) are coordinated, community-driven initiatives dedicated to improving research rigor, transparency, and methodological robustness. They operate through national subgroups (such as the Italian Reproducibility Network, IT-RN; the UK Reproducibility Network, UK-RN; and so on), and similar networks across Europe and beyond, which tailor activities to local needs while contributing to an international framework. RNs provide training, create educational materials (e.g., guides on preregistration and Registered Reports), and facilitate dialogue between researchers and institutions. A key function is offering a supportive environment for researchers adopting open practices, helping counteract feelings of professional isolation or increased scrutiny. Ultimately, RNs aim to foster a research culture grounded in transparency and trustworthiness.

The **Society for the Improvement of Psychological Science** (SIPS; https://improvingpsych.org/) represents a global, researcher-led community focused on advancing scientific practices in psychology. SIPS works through collaborative working groups, annual meetings, and community-driven projects that generate practical tools, guidelines, and consensus documents. Its mission overlaps with that of RNs, but SIPS functions more like a professional society, emphasizing interdisciplinary collaboration and the co-creation of solutions to methodological challenges.

The **Center for Open Science** (COS; https://www.cos.io/) complements these initiatives by providing both infrastructure and policy leadership for open research across disciplines. COS maintains the Open Science Framework (OSF), one of the most widely used platforms for preregistration, data sharing, and workflow transparency. In addition to offering extensive training and documentation, COS leads large-scale reproducibility efforts and collaborates with journals, funders, and institutions to implement open science standards.

Together, Reproducibility Networks, SIPS, and COS form a highly complementary ecosystem. While they differ in structure, they share core objectives: strengthening methodological rigor, promoting transparency, developing practical resources, and supporting researchers in adopting credible and reproducible scientific practices.

A checklist for transparent research

This checklist summarizes the key steps for ensuring statistical rigor and maximizing transparency in reporting, including aspects outlined, but not limited to, preregistration protocols. Please don't be overwhelmed by the level of detail in this checklist. It represents an ideal framework for planning, preregistration, and transparent reporting, but not every item will apply to every study. Think of it as guidance rather than a strict requirement. Every research project is unique, and sometimes, practical or ethical limitations (e.g., secondary data analysis, resource constraints) mean that not every item can be fully applied. Apply what is relevant to your research, and strive to follow these best practices as fully as possible within the limits of your design, resources, and context

Phase I: Planning

Component	Key Details to Include
1. Hypothesis specification	• Clearly label hypotheses (Primary vs. Secondary). • Specify direction of effect. • Include predictions (e.g., Cohen's d, η^2). • State falsification criteria.
2. Variable definition	• Define variables and instruments. • Specify context (e.g., lab setting). • Detail composite methods and transformations.
3. Sampling plan & stopping rule	• Rationale (e.g., power analysis). • Stopping rules and interim analyses.
4. Analysis & inference criteria	• Statistical models and inference criteria. • Multiple comparison corrections. • Directional vs. non-directional tests.
5. Contingency management	• Missing data and assumption testing. • Outlier and anomaly handling. • Sensitivity analyses.
6. Outcome-neutral checks	• Manipulation/fidelity checks. • Data quality inclusion/exclusion criteria. • Analysis plan for checks.

1. Essential Components of a Detailed Preregistration Plan

Phase II: Reporting and Disclosure

Component	Key Details to Include
7. Sharing materials	• Provide direct DOIs for preregistration, data, and code. • Deposit all materials (stimuli, instructions) in a public archive. • Specify file structure (e.g., "raw_data/", "codebook.pdf"). • Detail computational environment (software versions, packages). • Implement versioning for updates.
8. Disclose deviations	• Report all protocol deviations in a dedicated section. • Include methodological changes (e.g., new exclusion cutoffs). • Provide description, rationale, stage of occurrence, and impact. • View disclosure as transparency, not failure.
9. Analyses	• Detail software and functions for reproducibility. • Use consistent labeling to distinguish Confirmatory (preregistered) from Exploratory (post-hoc) results in all text and figures.
10. Data documentation & availability	• Share both raw and processed data. • Provide a codebook (structure, units, metadata). • Define missing values and variable coding. • Include a Data Availability Statement (DAS) with license info.
Additional: Conflict of interest	• Include a clear statement of financial, professional, or personal conflicts in the preregistration document.

2. Reporting, Transparency, and Open Science Practices

References

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The apa publications and communications board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Bakker, M., Veldkamp, C. L. S., van Assen, M. A. L. M., Crompvoets, E., Ong, H. H., Nosek, B. A., et al. (2020). Ensuring the quality and specificity of preregistrations. *PLOS Biology*, *18*(12), e3000937.

Bakker, M., Veldkamp, C. L. S., van den Akker, O. R., van Assen, M. A. L. M., Crompvoets, E., Ong, H. H., Nosek, B. A., et al. (2020). Recommendations in pre-registrations and internal review board proposals promote formal power analyses but do not increase sample size. *PLOS ONE*, *15*(7), e0236079.

Boeckhout, M., Zielhuis, G. A., & Bredenoord, A. L. (2018). The fair guiding principles for data stewardship: Fair enough? *European Journal of Human Genetics*, *26*(7), 931–936.

Chambers, C. D. (2013). Registered reports: A new publishing initiative at cortex. *Cortex*, *49*(3), 609–610. https://doi.org/10.1016/j.cortex.2012.12.016

Chambers, C. D., & Tzavella, L. (2022). The past, present, and future of registered reports. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-021-01193-7

Claesen, A., Gomes, S., Tuerlinckx, F., & Vanpaemel, W. (2021). Comparing dream to reality: An assessment of adherence of the first generation of preregistered studies. *Royal Society Open Science*, *8*, 211037.

Dunning, A., de Smaele, M., & Böhmer, J. (2017). Are the fair data principles fair? *International Journal of Digital Curation*, *12*(2), 177–195.

Frings, C. (2021). Experimental psychology joins peer community in registered reports. *Experimental Psychology*, *68*(1), 1–4.

Gernsbacher, M. A. (2018). Three ways to make replication mainstream. *Behavioral and Brain Sciences*, *41*, e129. https://doi.org/10.1017/S0140525X1800064X

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "estimating the reproducibility of psychological science." *Science*, *351*(6277), 1037. https://doi.org/10.1126/science.aad7243

Hasnain, A., & Rebholz-Schuhmann, D. (2018). Assessing fair data principles against the 5-star open data principles. In A. Gangemi et al. (Eds.), *The semantic web: Eswc 2018 satellite events* (pp. 574–580). Springer.

Heers, M. (2020). *Pre-registration and registered reports (fors guide no. 09, version 1.0)*. FORS. https://doi.org/10.24449/FG-2020-00009

Henderson, L. (2022). *A guide to preregistration and registered reports*. Preprint.

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., Courtot, M., Crosas, M., Dumontier, M., Evelo, C. T., Willighagen, E. L., Wittenburg, P., Roos, M., Mons, B., & Schultes, E. (2020). Fair principles: Interpretations and implementation considerations. *Data Intelligence*, *2*(1-2), 10–29. https://doi.org/10.1162/dint_r_00024

Kathawalla, U.-K., Silverstein, P., & Syed, M. (2021). Easing into open science: A guide for graduate students and their advisors. *Collabra: Psychology*, *7*(1), 18684. https://doi.org/10.1525/collabra.18684

Kush, R. D., Warzel, D., Kush, M. A., Sherman, A., Navarro, E. A., Fitzmartin, R., Pétavy, F., Galvez, J., Becnel, L. B., Zhou, F. L., Harmon, N., Jauregui, B., Jackson, T., & Hudson, L. (2020). Fair data sharing: The roles of common data elements and harmonization. *Journal of Biomedical Informatics*, *107*, 103421. https://doi.org/10.1016/j.jbi.2020.103421

Levitt, H. M., Motulsky, S. L., Wertz, F. J., Morrow, S. L., & Ponterotto, J. G. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology. *American Psychologist*, *73*(1), 26–46.

Martone, M. E. (2024). The past, present and future of neuroscience data sharing: A perspective on the state of practices and infrastructure for fair. *Frontiers in Neuroinformatics*, *17*, 1276407. https://doi.org/10.3389/fninf.2023.1276407

Martone, M. E., Garcia-Castro, A., & VandenBos, G. R. (2018). Data sharing in psychology. *American Psychologist*, *73*(2), 111–125. https://doi.org/10.1037/amp0000242

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.

Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly fair; revisiting the fair data guiding principles for the european open science cloud. *Information Services & Use*, *37*(1), 49–56. https://doi.org/10.3233/ISU-170824

Moshontz, H. (2018). *Licensing your work on psyarxiv*. Retrieved December 30, 2025, from https://blog.psyarxiv.com/2018/05/14/licensing-work-psyarxiv/

Moshontz, H., Binion, G. E., Walton, H., Syed, M., & Brown, B. T. (2020). *A guide to self-archiving preprints*. https://doi.org/10.31234/osf.io/dp4x9

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., et al. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A. (2014). Improving my lab, my science with the open science framework. *APS Observer*, *27*(3).

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Ofosu, G. K., & Posner, D. N. (2021). Pre-analysis plans: An early stocktaking. *Perspectives on Politics*, 1–17. https://doi.org/10.1017/S1537592721000931

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

Open Science Collaboration. (2016). Reproducibility and replicability in psychological science. *Science*, *351*(6277), 1037.

Pennington, C. R., & Heim, D. (2022). Reshaping the publication process: Addiction research & theory joins peer community in registered reports. *Addiction Research & Theory*, *30*(1), 1–4. https://doi.org/10.1080/16066359.2021.1931142

Sadeh, Y., Denejkina, A., Karyotaki, E., Lenferink, L. I. M., & Kassam-Adams, N. (2023). Opportunities for improving data sharing and fair data practices to advance global mental health. *Cambridge Prisms: Global Mental Health*, *10*, e14. https://doi.org/10.1017/gmh.2023.7

Sarafoglou, A., Kovacs, M., Bakos, B., Wagenmakers, E.-J., & Aczel, B. (2022). A survey on how preregistration affects the research workflow: Better science but more work. *Royal Society Open Science*, *9*(7), 211997. https://doi.org/10.1098/rsos.211997

Stewart, L., Rinke, E. M., McGarrigle, R., Lynott, D., Lautarescu, A., Galizzi, M., et al. (2020). *Pre-registration and registered reports: A primer from ukrn*. Retrieved December 30, 2025, from https://www.ukrn.org/primers/

Wagenmakers, E.-J., Sarafoglou, A., Aarts, S., Albers, C., Algermissen, J., Bahník, Š., van Dongen, N., Hoekstra, R., Moreau, D., van Ravenzwaaij, D., Sluga, A., Stanke, F., Tendeiro, J., & Aczel, B. (2021). Seven steps toward more transparency in statistical practice. *Nature Human Behaviour*, *5*(11), 1473–1480. https://doi.org/10.1038/s41562-021-01211-8

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The fair guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18

Zandonella Callegher, C., & Massidda, D. (2022). *The open science manual: Make your scientific research accessible and reproducible*. https://doi.org/10.5281/zenodo.6521850