



Reflections on Preregistration: Core Criteria, Badges, Complementary Workflows

Robert T. Thibault^{1,2,3}, Charlotte R. Pennington⁴,
Marcus R. Munafò^{2,3}

¹Meta-Research Innovation Center at Stanford (METRICS), Stanford University.

²School of Psychological Science, University of Bristol.

³MRC Integrative Epidemiology Unit at the University of Bristol.

⁴School of Psychology, Aston University.

Part of Special Issue

Reflections on the Unintended Consequences of the Science Reform Movement

Received

November 17, 2022

Accepted

April 12, 2023

Published

May 15, 2023

Issued

April 30, 2024

Correspondence

Stanford University
robert.thibault@stanford.edu

Funding

Robert T. Thibault is supported by a general support grant awarded to METRICS from Arnold Ventures and a postdoctoral fellowship from the Canadian Institutes of Health Research. Marcus Munafò and Robert Thibault are part of the MRC Integrative Epidemiology Unit (MC_UU_00011/7). The funders have no role in the preparation of this manuscript or the decision to publish.

License

This article is licensed under the **Creative Commons Attribution 4.0 (CC-BY 4.0)** license, which allows you to copy, redistribute, transform and build upon the article and all its contents in any medium or format for any purpose, provided that appropriate credit is given.

© Thibault, Pennington, & Munafò 2023



Clinical trials are routinely preregistered. In psychology and the social sciences, however, only a small percentage of studies are preregistered, and those preregistrations often contain ambiguities. As advocates strive for broader uptake and effective use of preregistration, they can benefit from drawing on the experience of preregistration in clinical trials and adapting some of those successes to the psychology and social sciences context. We recommend that individuals and organizations who promote preregistration: (1) Establish core preregistration criteria required to consider a preregistration complete; (2) Award preregistered badges only to articles that meet the badge criteria; and (3) Leverage complementary workflows that provide a similar function as preregistration.

Keywords *badges, blind data analysis, Open Science, prospective registration, preregistration*

Clinical trials are routinely preregistered¹ (Al-Durra2020). However, in other fields such as psychology and the social sciences only a small percentage of studies are preregistered,² and they often con-

tain ambiguities in the description of their study design, hypotheses, and analysis plans (Bakker2020vandenAkke2022). As advocates strive for broader uptake and more effective use of preregistration, the research community could benefit from drawing on the success of preregistration in clinical trials, where preregistration is commonplace.³ Preregistered clinical trials contain itemized and relatively explicit outcome measures, and most report their results.⁴

OSF instructs users to create a new preregistration). The prevalence of preregistration in psychology and the social sciences likely remains far below that of clinical trials research.

³For example, 92% of UK clinical trials submitting a final report to the NHS Health Research Authority in Sep 2021 to Sep 2022 are registered in a publicly accessible database (Authority2021); 75% of German clinical trials that were registered in 2017 were done so prospectively (BIHQUEST2023), and 71% of clinical trials published in PubMed-indexed journals in 2018 were registered, although only 42% were registered before the study began (i.e., preregistered) (Al-Durra2020).

⁴As of March 2023, for clinical trials that were completed over 12 months ago, 76% of those covered under the Food and Drug Administration Amendments Act and 84% of those posted on the European Union Clinical Trials Register (EUCTR) contained a link in their preregistration point-

¹Clinical trials research uses the term *prospective registration*, whereas other disciplines (including psychology and the social sciences) use the term *preregistration*. Prospective registration of clinical trials differs from preregistration of research in other disciplines in terms of its history, the functions it was designed to serve, and implementation details (e.g., clinical trial registries do not include sections specifically dedicated to hypotheses or analysis plans) (TARG2022). To streamline the reading of this commentary, we use the term *preregistration* for both clinical trials and other disciplines.

²Randomly sampled publications from the years 2014-2017 found that 1-5% (95% confidence intervals) of studies in psychology and 0-1% (95% CI) of studies in the social sciences were preregistered (Hardwicke2020Hardwicke2021). Another study found that 5% of over 90,000 articles in political science and international relations published from 2010-2021 were preregistered, but that 16% were preregistered in 2021 (Scoggins2023). Other sources also indicate an increase in the prevalence of preregistration since Hardwicke et al.'s 2014-2017 sample. For example, there were 12,000 OSF preregistrations from 2012-2017 (Nosek2018b), but over 100,000 across a similar time period from 2018 to January 2023. Some of these preregistrations, however, may be duplicates (e.g., to update a preregistration, the

We propose three actions for the research community to consider to improve the function of preregistration in psychology and the social sciences (see Table 1 & Box 1). These proposals stem from insights developed while conducting research on preregistration across disciplines, including meta-analyses of discrepancies between preregistrations and published manuscripts (TARG2021) and a feasibility study of a peer review intervention to address these discrepancies before publication (TARG2022). We discuss the function of preregistration in terms of reducing bias and making risk of bias transparent (Hardwicke2023), as well as the auxiliary benefit of improved research quality.⁵ Our proposals are by no means exhaustive; more comprehensive overviews of preregistration are available elsewhere (Hardwicke2023DeVito2022).

We propose that advocates for preregistration consider to:

1. **Establish core preregistration criteria** (i.e., a minimum amount of information required to consider a preregistration complete—as the World Health Organization’s International Clinical Trials Registry Platform has done for clinical trial registration).
2. **Award preregistered badges only to articles that meet the badge criteria** (of which few currently do).
3. **Leverage complementary workflows that provide a similar function as preregistration** (e.g., blinded data analysis to minimize data-dependent analytical decisions).

1. Establish core preregistration criteria

For clinical trials to be considered fully registered, they must provide information regarding 24 specific items, known as the Trial Registration Data Set (Organization2017). Although these 24 items do not include a detailed analysis plan, they set a minimal standard that organizations such as the International Committee of Medical Journal Editors can promote

ing to the reported results. Notably, however, reported outcomes do not always match preregistered outcomes (see footnote 7).

⁵We acknowledge that there are ongoing debates about the function, purpose, and goals of preregistration (McPhetres2020). We feel that this debate does not preclude implementation of our suggestions.

(ICMJE2022ICMJE2023). This itemized standard laid the foundation for regulations and institutional infrastructure, which in turn drove the widespread uptake of preregistration in clinical trials.⁶ It allows for transparent updating of preregistrations and makes comparisons between preregistrations and publications relatively easy (see Figure 1). The structure is sufficiently clear-cut, such that the Health Research Authority in the UK now uses information from ethics applications to register trials on behalf of clinical trialists (Authority2021). These researchers can go beyond the minimum 24 items, add as many details as they would like to the registration, and append a study protocol.⁷

In contrast to clinical trial registrations which include discrete items followed by short responses (e.g., primary outcome; sample size), preregistration templates in psychology and the social sciences often include broad headers followed by blocks of text (e.g., hypotheses, analysis plan—see Figure 1).⁸ Single hypotheses can contain multiple elements that would be better divided into several distinct hypotheses. Preregistrations sometimes list several variables and analyses but provide a sample size calculation for only one analysis. Within a preregistration, aligning a single hypothesis to its outcome measure and analysis can be far from trivial. Matching these to text in the manuscript presents an additional challenge.

⁶For example, in 2005, the International Committee of Medical Journal Editors (ICMJE) made preregistration a necessary condition for consideration for publication (DeAngelis2004). Clinical trial registration rates quadrupled in 2005 and thousands of journals now claim to follow this policy (ICMJE2023).

⁷Notably, clinical trial preregistration is not a panacea. Some trials are never preregistered; among those that are, about one-third publish at least one different primary outcome than what was preregistered, and about two-thirds publish at least one different secondary outcome than what was preregistered (TARG2021). While some researchers challenge the usefulness of clinical trial preregistration based on this situation (Lash2022Abrams2020), it is trial preregistration that allows us to identify this risk of bias. Without preregistration, we would not be able to quantify the level of selective reporting or publication bias. Moreover, preregistering protocols—which include statistical analysis plans—remains rare in biomedicine broadly (Serghiou2021).

⁸Granted, the scope of preregistration must be expanded to capture the breadth of study designs used in psychology and the social sciences, as compared to the more limited design options used in clinical trials.

Table 1 Problems and proposed solutions for preregistration in psychology and the social sciences.

Problem	Proposed solutions*
<ol style="list-style-type: none"> 1. Low uptake of preregistration 2. Imprecise and ambiguous preregistrations 3. Poor alignment between preregistrations and manuscripts 	<ol style="list-style-type: none"> A. Establish core preregistration criteria B. Award preregistered badges only to articles that meet the badge criteria C. Leverage complementary workflows that provide a similar function as preregistration

1. Preregistration provides additional benefits once a substantial proportion of studies are preregistered—such as facilitating evidence synthesis and reducing duplication.
2. Imprecise language and ambiguities in preregistrations leave them open to various interpretations, and in turn, limit their ability to reduce bias and transparently communicate study plans.
3. Poor alignment between preregistrations and manuscripts—both in terms of the overall structure of the documents as well as the specific content—make it difficult to compare the texts to assess risk of bias.

The solutions we propose are partial solutions in the sense that they remain unlikely to fully solve shortcomings in preregistration. They can be implemented individually or alongside other efforts.

*We itemize the *problems* with numbers and the *proposed solutions* with letters to indicate that they are not aligned in a one-to-one manner; each proposed solution could impact each of the three problems to different extents.

Thus, the less structured information provided in many psychology preregistrations can obscure a reader's understanding of what the researchers planned to do and whether they did it.^{9,10}

Given the low prevalence of preregistration in psychology and social sciences research (Hardwicke2020Hardwicke2021Scoggins2023) alongside the difficulty of comparing preregistered study details to published study reports (TARG2022vandenAkker2022) we argue that establishing core preregistration criteria would complement ongoing initiatives that strive for ideal practice.¹¹

⁹We have no qualms when researchers deviate from a preregistration, so long as they disclose the deviation when reporting results. In many cases, deviations are entirely justifiable and can be necessary to improve a study design or analysis.

¹⁰One complication for preregistration in psychology and the social sciences, as compared to clinical trials, is that the spectrum of research questions ranges from highly exploratory to purely confirmatory. Some ambiguities in preregistrations may arise when researchers attempt to present exploratory research questions in a confirmatory format.

¹¹Some researchers propose that publications of preregistered studies should include a section outlining deviations

from the preregistration (Campbell2019). While we agree with these initiatives, sections on deviations will be difficult to interpret if preregistrations are highly ambiguous due to a lack of itemization.

Efforts have been made to create standard preregistration templates in psychological science (e.g., Open Science Framework, AsPredicted), but these can vary substantially, and there is no broad agreement regarding the details they must include. In an attempt toward standardization, a Preregistration Task Force consisting of the American Psychological Association (APA), the British Psychological Society (BPS), and the German Psychological Society (DGP), supported by the Center for Open Science (COS) and the Leibniz Institute for Psychology, developed a consensus template¹² for the preregistration of quantitative psychology research (Bosnjak2022). On the one hand, the template is exhaustive and was purposefully designed to parallel the structure of the APA Style Journal Article Reporting Standards (Appelbaum2018); its proper use

from the preregistration (Campbell2019). While we agree with these initiatives, sections on deviations will be difficult to interpret if preregistrations are highly ambiguous due to a lack of itemization.

¹²Although the authors use the terms *consensus* and *consensus template* throughout their manuscript, there is no description of the consensus process, which may not have been formalized.

would present a very effective implementation of preregistration. On the other hand, there is no evidence that user-testing informed the template¹³ and its uptake remains limited at this time.¹⁴

Comparable guidelines in clinical trials have been developed through formal consensus processes that involve diverse stakeholders, include a user-testing stage (i.e., piloting), and are widely used by researchers.¹⁵ These documents were designed to apply across clinical trials research, regardless of the specific discipline. Their structure is such that researchers can create extensions to the guidelines to target their specific disciplines more fully (Zhang2020Thabane2016). Core preregistration criteria could be developed through a similar process and designed to accommodate the diversity of study types in psychology and the social sciences. They could facilitate broad adoption of preregistration by setting a minimum standard that is relatively easy to achieve and a benchmark upon which publishers, funders, and institutions can develop regulations.

2. Award preregistered badges only to articles that meet the badge criteria

As of April 2023, the Center for Open Science website lists 80 journals that award badges to articles that claim to have used open science practices such as preregistration, open materials, and open data (www.cos.io/initiatives/badges). To receive a preregistered badge, a publication should have no undisclosed discrepancies from the preregistration (COS2023). And yet, two studies analyzing psychology publications with preregistered badges found that

89% of 27 articles contained at least one undisclosed discrepancy (Claesen2021)¹⁶ and 67% of 258 articles selectively reported at least one hypothesis (vandenAkker2022).¹⁷ The organization that developed the badges—the Center for Open Science—describes two ways to award badges: author self-disclosure or peer review (COS2016).

Disclosure

Some journals, in their instructions for authors, state that they use the self-disclosure method to award badges (e.g., *Psychological Science*, *Journal of Experimental Social Psychology*), but the disclosure statement provided by the COS which these journals use does not align with the preregistered badge criteria. The four criteria for a preregistered badge are: “(1) A public date-time stamped registration is in an institutional registration system; (2) Registration pre-dates the intervention; (3) Registered design corresponds directly to reported design; and (4) Full disclosure of results in accordance with registered plan” (COS2023). However, the disclosure form used by these journals asks authors to complete five disclosure items (COS2016), none of which match the third and fourth badge criteria. Thus, authors can both truthfully complete the disclosure form and not meet the badge criteria. Even if the disclosure items were realigned to match the badge criteria, it remains unclear whether the proportion of badged papers that fully meet all criteria would rise in the absence of a verification mechanism.

Peer review

We are not aware of any journal that systematically peer reviews articles to ensure they meet the criteria for a preregistered badge. Moreover, based on our experience (TARG2022) and that of other researchers who have systematically examined publications awarded with a preregistered badge (Olmo van den Akker, personal communication, 2021; Aline

¹³Two of the authors are now testing the usability of the PRP-QUANT template (Spitzer2021).

¹⁴This manuscript was first posted as a preprint in February 2021 and has received 27 citations (according to Google Scholar on 31 Jan 2023). These citations appear to mostly reference, rather than use, the PRP-QUANT template. If this template serves a broad user base, is user-friendly, and strongly supported by the organizations involved in making the template (some of the leading psychology organizations in the world), we would predict a greater uptake to date, though increased uptake may occur in the years to come. As of April 2023, it appears the template has been used up to 75 times on the PsychArchives repository.

¹⁵For example, the Consolidated Standards of Reporting Trials (CONSORT), the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) (Hopewell2022).

¹⁶Researchers replicated and extended this study and found comparable results (Weaver2022).

¹⁷This sample also includes studies without a preregistered badge, although most articles did have a preregistered badge. They sampled from two populations: articles that contained a preregistered badge and articles that earned a Preregistration Challenge Prize from the Center for Open Science. Their manuscript, however, only reports summary results across both samples.

Claesen, personal communication, 2019), we feel it is very difficult to confidently state that the “Registered design corresponds directly to reported design” or “Full disclosure of results in accordance with registered plan.” Indeed, one study found that researchers could only agree on the number of hypotheses present in 14% of preregistrations (**Bakker2020**). Another study with a strict operationalization of what constituted a hypothesis had 54% agreement between coders regarding the number of hypotheses (**vandenAkke2022**). The lack of itemized core preregistration criteria alongside differences in the structure of preregistrations and manuscripts renders many comparisons ambiguous.

One could argue that the issues we present regarding inaccurate awards are outweighed by the benefit that badges may have on the uptake of preregistration. Indeed, the badge criteria were designed to represent a high aspirational standard, rather than setting a minimal bar. However, there is a possibility that badges in their current implementation have negative effects. Given that most articles awarded a preregistered badge do not fully meet the criteria for earning that badge, awarding badges can create a false impression that rigorous research practices are being used and therefore lend undue trust to studies awarded a preregistered badge. This practice could also have downstream impacts on the trustworthiness of these types of initiatives more broadly.¹⁸

Changing the criteria for the preregistered badge could be one way to make clearer what the badge signals. They could be revised, for example, to require only the existence of a permanent and public preregistration in a repository that provides a DOI, without requiring that the preregistration was followed. This criterion would be easy to audit¹⁹ and achieves at least one main function that preregistration was designed to address—putting a timestamp on study plans to help demarcate confirmatory research (**Nosek2018**). An additional criterion

could demand that the preregistration include all the items in an established core preregistration criteria, as outlined earlier in this commentary. Machine-readable preregistration statements could also be employed to facilitate automated compliance monitoring from funders or institutions. Based on the current badge criteria, a publication whose preregistration has almost no detail could earn a preregistered badge, whereas a publication with a very detailed preregistration and a minor discrepancy should not earn a badge.²⁰

Taken together, current practices for awarding preregistered badges reward researchers even if their preregistration is of low quality and aligns poorly with the associated publication. We commend the development and testing of new initiatives; at the same time, we advocate for follow-up and evaluation to investigate whether they work as intended.

3. Leverage complementary workflows that provide a similar function as preregistration

Preregistration can reduce bias, increase transparency, and may also improve research quality (**Hardwicke2023Sarafoglou2022b**). However, there are no checks and balances to evaluate whether the study outlined in a preregistration is well designed or clearly described (**Chambers2022**). Journal policies and peer review can improve the quality of reporting in relation to a preregistration, but they occur too late in the research pipeline to impact the study design or preregistration quality. Complementary research workflows could achieve some of the same functions as preregistration and may come with additional benefits (e.g., blind data analysis, Experimental Design Assistants, protocol peer review).²¹

For observational research, data management organizations could employ workflows that necessitate open research practices. For example, they could provide researchers with a synthetic dataset, which researchers could

¹⁸For example, as has happened for “data available upon request” statements, where over 90% of requests go unanswered or are declined (**Gabelica2022**).

¹⁹Auditing could be conducted by any one of various stakeholders—for example, by the journals who reward the badges, the organization who developed them (COS), funders, researchers (**DeVito2019**), or institutions (e.g., similar to the Charité Dashboard on Responsible Research).

²⁰Although in practice, badges appear to be awarded in both cases. When appropriately disclosed, we have no qualms with discrepancies.

²¹See **Srivastava2018** for a more in-depth discussion of various alternatives and complementary strategies to typical preregistration.

use to develop an analysis script. The researchers would then run their analysis in a Trusted Research Environment (TRE) where the results are output, the real data remains hidden, and the analysis is logged and made public (e.g., as done at OpenSAFELY.org). If a Trusted Research Environment is not available, data management organizations could simply provide the complete dataset after the researchers register their analysis script (Sarafoglou2022Thibault2023). These workflows make executable analyses—as opposed to sometimes ambiguous blocks of text—publicly available, while also protecting researchers from making data-dependent analytical decisions. They also overcome arguments raised against preregistration for observational research, including that the data often already exist, knowledge of the data may be necessary to devise a reasonable analysis plan, and registration can inhibit exploration (Lash2012).

Researchers can execute a comparable workflow for experimental studies by writing an analysis script based on simulated data and preregistering it before beginning data collection. In other words, the preregistration would include a results section based on a simulated dataset and the numbers are simply updated after running the analysis on the real data.

Another research tool—the Experimental Design Assistant—can be employed in a similar manner. This web application, developed by The National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), uses a graphical interface to walk pre-clinical animal researchers through designing their experiment. Unlike pEDA is an interactive tool that requires the user to input thorough information when the input fails to align with the PDF which the NC3Rs encourage to append to funding applications (NC3Rs2021). The tool has the potential to simultaneously reduce the design of effective experimental workflows by using the PDF as a component of a preregistration.

These examples hold the potential to increase the uptake of preregistration, improve the precision of preregistration, and do so by embedding the research into a user-friendly workflow that facilitates study plans.

Conclusion

In psychology and the social sciences, preregistrations can reduce bias and improve transparency. At the same time, they remain under-used, can lack clarity, and are often difficult to compare directly with their associated publication. Current efforts to promote the uptake of preregistration (e.g., badges) and improve preregistration quality (e.g., the PRP-QUANT template) rely largely on the willingness and scrupulousness of research teams alone. We propose that the research community consider parallel initiatives to simplify and standardize preregistration (e.g., adopt itemized core preregistration criteria), and to leverage complementary workflows that necessitate open research practices.

Contributions

RTT wrote an initial draft. All other authors were involved in relevant discussions and contributed to the final draft.

Funding

Robert Thibault is supported by a general support grant awarded to METRICS from Arnold Ventures and a postdoctoral fellowship from the Canadian Institutes of Health Research. Marcus Munafò and Robert Thibault are part of the MRC Integrative Epidemiology Unit (MC_UU_00011/7). The funders have no role in the preparation of this manuscript or the decision to publish.

Box 1. Agents of change

Our three recommendations target the research community broadly and differ in their implementation pathways. Establishing core preregistration criteria would require coordination across various stakeholders including publishers, funders, institutions, learned societies, researchers, and other end-users of research findings. Such an initiative would take a concerted effort and could gain momentum through a grassroots push from researchers or a top-down mechanism from major funders. In contrast, any journal can improve their own use of preregistered badges, and the organization who created them—the Center for Open Science (COS)—holds the ability to redefine the badge criteria. As for our final recommendation, any individual research group, funder, data management organization, or other stakeholder can explore the use of complementary workflows. Highly successful workflows could then be adopted more widely.

| Acknowledgements

We thank Gustav Nilsson, Steven Goodman, Mario Malički, Marton Kovacs, and Lisa Spitzer for feedback on earlier drafts of this commentary.

| Competing interests

All other authors declare no conflict of interest.

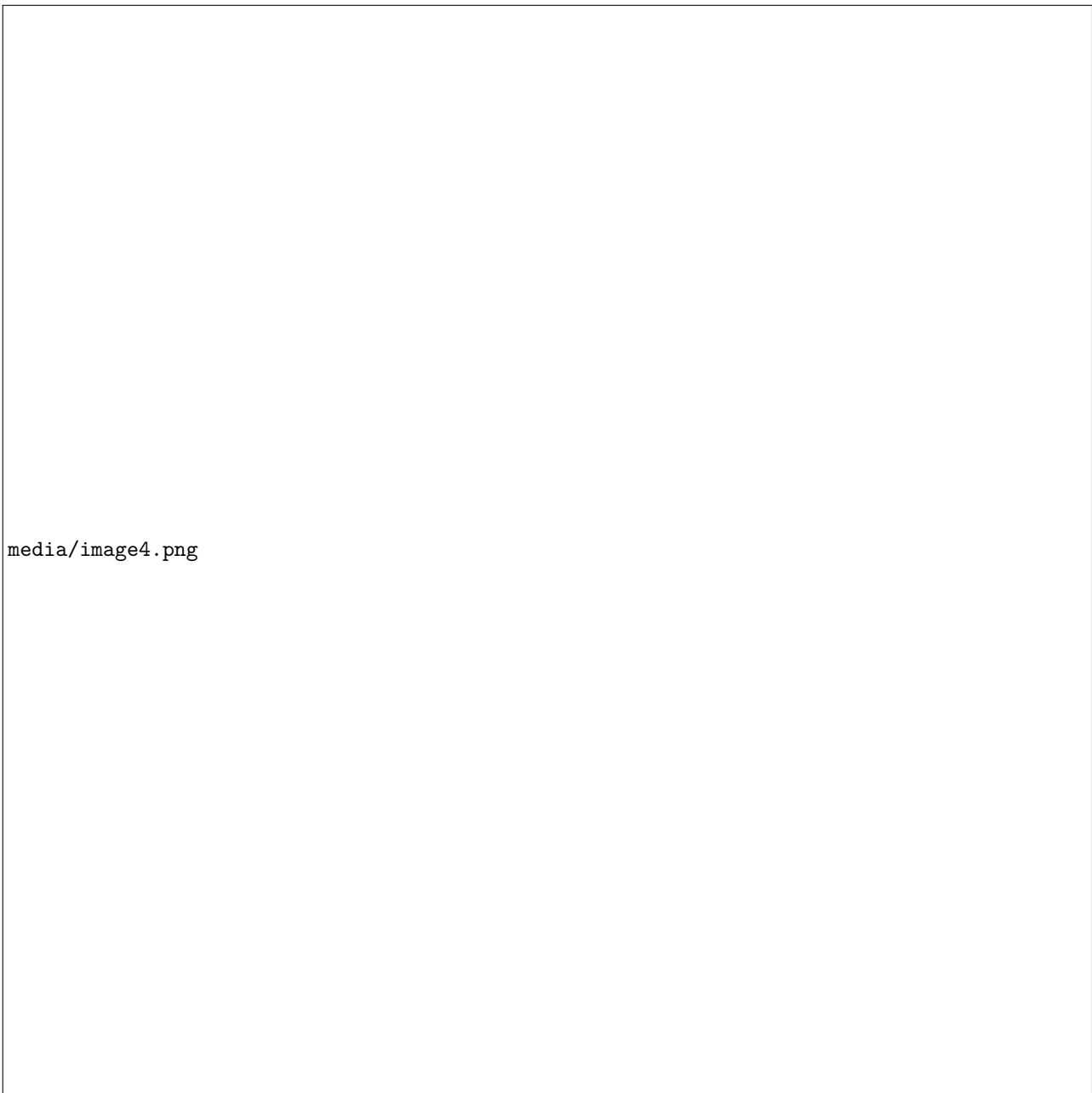


Figure 1 Comparison of a clinical trial preregistration excerpt (left) to an OSF preregistration excerpt (right)*.

The clinical trial preregistration excerpt demonstrates several features that the psychology and social sciences community could benefit from considering. These include: (1) The option for an itemized and tabular format; (2) Clear demarcation of the 24 items contained in the WHO Trial Registration Data Set, which is supported by the International Committee of Medical Journal Editors and demarcated with the superscript ^{ICMJE}; (3) Clear demarcation of the primary outcome measure and time frame of assessment; (4) Easy identification of updates to the primary outcome measure (e.g., several time points were added); (5) Easy identification of core items that are not provided (e.g., secondary outcome measures); and (6) A link to a *Change History* log, which looks similar to a Microsoft Word document with track changes. The OSF has recently begun to provide a function allowing researchers to update their preregistration. The updated preregistration identifies sections that were updated (e.g., “Hypotheses”), but does not provide the track-changes style functionality that clinicaltrials.gov does. OSF preregistrations also often contain a statistical analysis plan whereas clinical trial preregistrations rarely do.

* These excerpts are copied from (Arnold2013) and (Berent2021). We selected the clinical trials registration because the lead author (RTT) was familiar with it, and it clearly depicts several benefits of clinical trial preregistrations in a relatively small screenshot. The OSF registration

was selected by going to www.osf.io/registries and selecting Provider: “OSF Registries” and OSF Registration Type: “OSF Preregistration,” and then choosing a recent preregistration that depicts the text-block response format.

Thibault, Pennington, & Munafò (2023). Reflections on Preregistration: Core Criteria, Badges, Complementary Workflows. *Journal of Trial & Error*, 4(1), 20–21. <https://doi.org/10.36850/mr6>.