

# Preregistration of Psychology Meta-Analyses: A Cross-Sectional Study of Prevalence and Practice



Alejandro Sandoval-Lentisco<sup>1</sup>, Miriam Tortajada<sup>1</sup>,  
Rubén López-Nicolás<sup>1</sup>, José A. López-López<sup>1</sup>, E. J. Wagenmakers<sup>2</sup>,  
Julio Sánchez-Meca<sup>1</sup>, and Tom E. Hardwicke<sup>3</sup>

<sup>1</sup>Department of Basic Psychology and Methodology, University of Murcia, Murcia, Spain; <sup>2</sup>Department of Psychology, University of Amsterdam, Amsterdam, Netherlands; and <sup>3</sup>Melbourne School of Psychological Sciences, University of Melbourne, Australia

Advances in Methods and Practices in Psychological Science  
January-March 2025, Vol. 8, No. 1,  
pp. 1–12  
© The Author(s) 2025  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/25152459241300113  
www.psychologicalscience.org/AMPPS



## Abstract

Meta-analyses play an influential role in synthesizing the existing evidence on a particular topic. Consequently, it is especially important that meta-analyses are conducted and reported to the highest standards and that the risk of bias is minimized. Preregistration can help detect and reduce bias arising from opportunistic use of “researcher degrees of freedom.” However, little is known about the prevalence and practice of the preregistration of meta-analyses in psychology. In this study, we first measured the prevalence of preregistration in all psychology meta-analyses published in 2021. Next, for 100 randomly selected preregistered meta-analyses, we evaluated the preregistration’s coverage of key meta-analytic decisions and the extent to which published meta-analyses deviated from their preregistered protocols. Of all 1,403 eligible psychology meta-analyses published in 2021, 382 (27%) were preregistered. In our random sample, we found that key PRISMA-P decision items were often omitted from preregistered protocols—out of the 23 decision items that were examined, the median number of items covered was 13 (interquartile range [IQR] = 11–14). We also found that all 100 preregistered meta-analyses contained at least one deviation from the preregistered protocol ( $Mdn = 9$ ,  $IQR = 6.75–11$ ) and that most deviations were undisclosed ( $Mdn = 8$ ,  $IQR = 6–11$ ). These findings suggest that the infrequent use and poor implementation of preregistration in psychology meta-analyses undermines its potential to reduce bias and increase transparency.

## Keywords

preregistration, meta-analysis, PRISMA-P, coverage, deviations

Received 5/10/24; Revision accepted 10/21/24

Conducting a meta-analysis requires researchers to make multiple decisions, such as determining which studies to include, specifying the outcomes of interest, and selecting the statistical model (Johnson, 2021; Moreau & Gamble, 2022; Pigott & Polanin, 2020). Each decision may have several justifiable options (“researcher degrees of freedom”; Simmons et al., 2011; Wicherts et al., 2016), creating a multiplicity of possible outcomes. When researchers make decisions after seeing the data (data-dependent decisions), this may introduce bias, particularly when those decisions are influenced by whether the results align with the researchers’ preferred outcome

(Hardwicke & Wagenmakers, 2023; Wicherts et al., 2016). Considering their influential role in synthesizing the existing evidence on a particular topic, it is especially important that meta-analyses are conducted and reported to the highest standards and that the risk of bias is minimized.

## Corresponding Author:

Alejandro Sandoval-Lentisco, Department of Basic Psychology and Methodology, University of Murcia, Murcia, Spain  
Email: alejandro.sandovall@um.es



Preregistering a study protocol has been proposed as a tool for detecting and reducing data-dependent biases arising from researcher degrees of freedom (Munafò et al., 2017). The degree to which a preregistration minimizes the risk of bias depends on how clear and comprehensive that preregistration is, that is, how well the preregistration constrains data-dependent decision-making (Bakker et al., 2020; Hardwicke & Wagenmakers, 2023). In addition, deviations from preregistered decisions can increase the risk of bias (even if they may have other benefits, such as improving the validity of a poorly planned analysis) and should be explicitly disclosed and justified. It is therefore important to evaluate how often meta-analyses are being preregistered (i.e., assess their prevalence) and how effective those preregistrations may be at minimizing the risk of bias (i.e., assess their coverage of researcher degrees of freedom and the frequency and nature of any deviations).

Previous studies assessing preregistration prevalence have found that preregistration of empirical studies (including meta-analyses) is infrequent in psychology and the social sciences (e.g., Hardwicke et al., 2020, 2022), although increasing (Hardwicke et al., in press). Some recent studies have specifically examined the prevalence of preregistered protocols for systematic reviews and meta-analyses in subdomains of psychology. Polanin et al. (2020) found that 2% of 150 meta-analyses published in *Psychological Bulletin* from 1990 to 2017 were preregistered. López-Nicolás et al. (2022) found that 19% of a random sample of 100 meta-analyses on clinical-psychological interventions published between 2000 and 2023 were preregistered. Sandoval-Lentisco et al. (2024) found that 36% of 97 meta-analyses on cognitive training published between 1997 and 2023 (of which 40 were published from 2020) were preregistered. Currently, there is no available estimate for the prevalence of preregistered meta-analyses across the entire field of psychology.

When empirical studies are preregistered, they often provide poor coverage of researcher degrees of freedom, such as when to stop data collection or how to define outliers (Bakker et al., 2020). In addition, deviations from preregistered protocols are frequent and seldom disclosed or discussed (Claesen et al., 2021; TARG Meta-Research Group & Collaborators, 2023; van den Akker et al., 2023). In the context of systematic reviews and meta-analyses, some studies have assessed protocol coverage, but these have mostly focused on the prespecification of outcomes of interest in medical trials. These studies found, for example, that 35% to 38.5% of assessed systematic reviews did not preregister any outcome measure (Parmelli et al., 2007; Tricco et al., 2016). Moreover, these and other studies suggest that deviations from preregistered protocols of systematic reviews and meta-analyses in medicine are common. Silagy et al. (2002)

found that 43 out of 66 Cochrane reviews contained at least one major deviation from their preregistered protocol, including changes in the objectives, outcome measures, methods, and inclusion criteria. Studies have also found that 22% to 39% of systematic reviews and meta-analyses omitted or changed the outcomes of interest compared with what was stated in the registered protocol (Dwan et al., 2013; Kirkham et al., 2010; Page et al., 2014; Tricco et al., 2016). To our knowledge, no study has examined preregistered protocol coverage and deviations for meta-analyses published in psychology.

## Aims of This Study

The present study had three aims. First, we measured the prevalence of preregistration in all meta-analyses published in psychology in 2021. Second, we assessed the coverage of preregistrations by examining the extent to which 100 randomly selected preregistered meta-analyses published in 2021 clearly prespecified key meta-analytic decisions that may incur a risk of bias. Finally, we assessed deviations from preregistrations by comparing the decisions reported in the final article with those specified in the preregistered protocol (for the same 100 randomly selected preregistered meta-analyses). When we identified deviations, we checked whether they were explicitly acknowledged and justified in the final article.

## Disclosures

This study was preregistered on September 1, 2022 (<https://osf.io/rnv5c>). After starting data collection, we made minor amendments to the inclusion criteria on October 18, 2022, and again on February 20, 2023 (see the different versions of the protocol at <https://osf.io/kf9dz>). All deviations from the original protocol are disclosed and explained in <https://osf.io/6kxgt>. All materials, primary data, and analysis code are available at <https://osf.io/ru5h4/>. Below, we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

## Method

### Sampling Strategy

For our first research aim, measuring prevalence, the population we wished to generalize to was all meta-analyses published in psychology journals in 2021. To identify this population, we searched for meta-analyses published in journals included in the category of psychology in 2021 in the Web of Science Core Collection using the following search string in the Web of Science search platform accessed through the University of

Murcia: ((TS = (meta-analy\*)) AND WC = (Psychology)) AND PY = (2021).

For the second and third research aims, measuring coverage and deviations, we also wished to generalize to all preregistered meta-analyses published in psychology journals in 2021. Because the assessment process for the second and third research aims entailed considerable manual workload, we examined 100 randomly sampled preregistered meta-analyses rather than the entire population. This quantity was selected because it was within our resource constraints and yielded estimates with a margin of error that we believe to be reasonably accurate—a precision analysis indicated that the margin of error on a 95% confidence interval (CI) for a proportion of 0.5 was 0.098 (see Appendix A of the preregistered protocol, <https://osf.io/rnv5c>). We performed the random selection by generating a random number for each meta-analysis that was associated with a preregistration, sorting the meta-analyses by their assigned number, and choosing the first 100 eligible articles.

### ***Inclusion Criteria***

For articles to be included, we had to have full-text access, and they had to be in English language. For all research aims, articles were considered eligible if they conducted a meta-analysis synthesizing the evidence on any topic using any kind of meta-analytical methodology. We defined *meta-analysis* as “the statistical combination of results from two or more separate studies,” according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks et al., 2023, p. 243). Metaresearch or methodological articles related to the development of meta-analytic methods or evaluation of meta-analyses in practice were not included. Reanalyses of meta-analyses or internal meta-analyses were not included. Systematic reviews without a statistical synthesis of effect sizes were also not included. After conducting the search and checking the eligibility of each record, we also identified different types of studies that we did not consider before commencing the study (see deviation at <https://osf.io/6kxgt>).

For the second and third research aims, articles were not included in the random sample if we detected that the time between the registration date and the submission date was less than 1 month apart. When such articles were identified, they were excluded and replaced with the next one in the list.

### ***Extraction and Coding of Information***

For our first research aim, measuring prevalence, we manually inspected the title and abstract of all the articles identified by our search to see if they were eligible

meta-analyses according to our inclusion criteria. If we had any doubts as to whether the article fulfilled our criteria, we inspected the full text. For eligible meta-analyses, we coded whether they reported having a preregistered protocol published in a journal or posted in any kind of repository/registry (e.g., PROSPERO, OSF, clinicaltrials.gov, or an institutional website) and whether it was accessible. For this inspection, we manually opened the document and performed a keyword search (ignoring case) with the words “registered,” “registration,” “protocol,” “PROSPERO,” “NCT,” and “OSF.” If searching for those keywords retrieved no results, we also manually inspected the abstract, author note, the start of the methods section, and any disclosure statements appearing after the discussion section. Two authors (A. Sandoval-Lentisco and M. Tortajada) independently coded all the articles. A third author (T. E. Hardwicke) helped to resolve coding differences. The interrater agreement of this process was high (96% agreement for assessing the eligibility; 98% agreement for assessing whether an article referred to a preregistration).

For the second and third research aims, assessing coverage and deviations, we extracted information about the research decisions that were reported in the preregistered protocols and their corresponding final articles using a structured coding form (see <https://osf.io/tgh2p>). For 62 articles, we found more than one protocol version available. In these cases, we based our assessment on the latest protocol version available before the data analysis, discarding those versions that were uploaded after the study was submitted to the journal or marked as completed.

We did not aim to cover all possible research decisions; given the diversity of meta-analytic methods, the number of possible decisions is likely to be enormous and difficult to specify. Instead, we decided to evaluate key decisions that are applicable to any meta-analysis as outlined in the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P), an established reporting guideline based on an expert-consensus development process (Moher et al., 2015).<sup>1</sup> PRISMA-P includes a wide range of research decisions, such as the reporting of eligibility criteria, outcomes of interest, data synthesis, or methods to assess biases.<sup>2</sup>

For most research decisions, we did not know the likelihood that a researcher would actually exploit that degree of freedom in practice (knowingly or unknowingly) or the potential impact on the results; we therefore included any decision noted in PRISMA-P that could potentially incur a risk of bias.<sup>3</sup> Specifically, we selected every item from PRISMA-P that refers to researcher decisions that can be adjusted after the results are known and can potentially change the results (see Supplemental Table 1 for the items that were included and any

modifications from the original PRISMA-P checklist and Supplemental Table 2 for the original PRISMA-P items that were not included at <https://osf.io/jq792>). For cases in which single PRISMA-P items covered several research decisions, we divided these into individual items accordingly.

For each decision item, we recorded whether it was specified in the protocol, whether it was later mentioned in the final article, whether there was any deviation, and in case we found any deviation, whether the deviation was disclosed and explained. We considered that there was a deviation when the substantive information regarding a decision item differed between the protocol and the final article. We did not consider minor grammatical differences, such as inconsequential changes in wording, to be deviations. We distinguished three types of deviations: (a) decision item present in the protocol but not in the final article (an “omission”), (b) decision item in the final article but not specified in the protocol (an “addition”), and (c) decision item specified differently in the protocol and final article (a “modification”).

The assessment of coverage and deviations was performed in duplicate: A primary coder (A. Sandoval-Lentisco) coded all the articles, and two secondary coders (M. Tortajada and R. López-Nicolás) independently coded 72 and 28 articles each, respectively. The independent coding and resolution of coding differences was performed, periodically, in rounds of four to 10 articles. Information regarding protocol versions and time between protocol registration and manuscript submission was extracted by a sole author (A. Sandoval-Lentisco). For checking whether information was present in the protocol and final article and whether there was any deviation, the percentage agreement between raters ranged from 76% to 100% (for more details, see Supplemental Table 3 at <https://osf.io/jq792>).

## Results

Below we report 95% CIs based on the Wilson score interval (Wilson, 1927) for binomial items and the Sison-Glaz method for multinomial items (Sison & Glaz, 1995). For information regarding the time between registration and publication, the number of protocol versions, and the research stages at which the studies were registered, see <https://osf.io/2ab8n>.

### Identification of Eligible Articles

Figure 1 summarizes the process of identifying eligible articles. Our search for meta-analyses published in psychology journals in 2021 identified 2,137 articles. Of these, 1,403 (66%) were considered eligible, 726 (34%) were considered not eligible, and we could not access the full text for eight (0.3%).

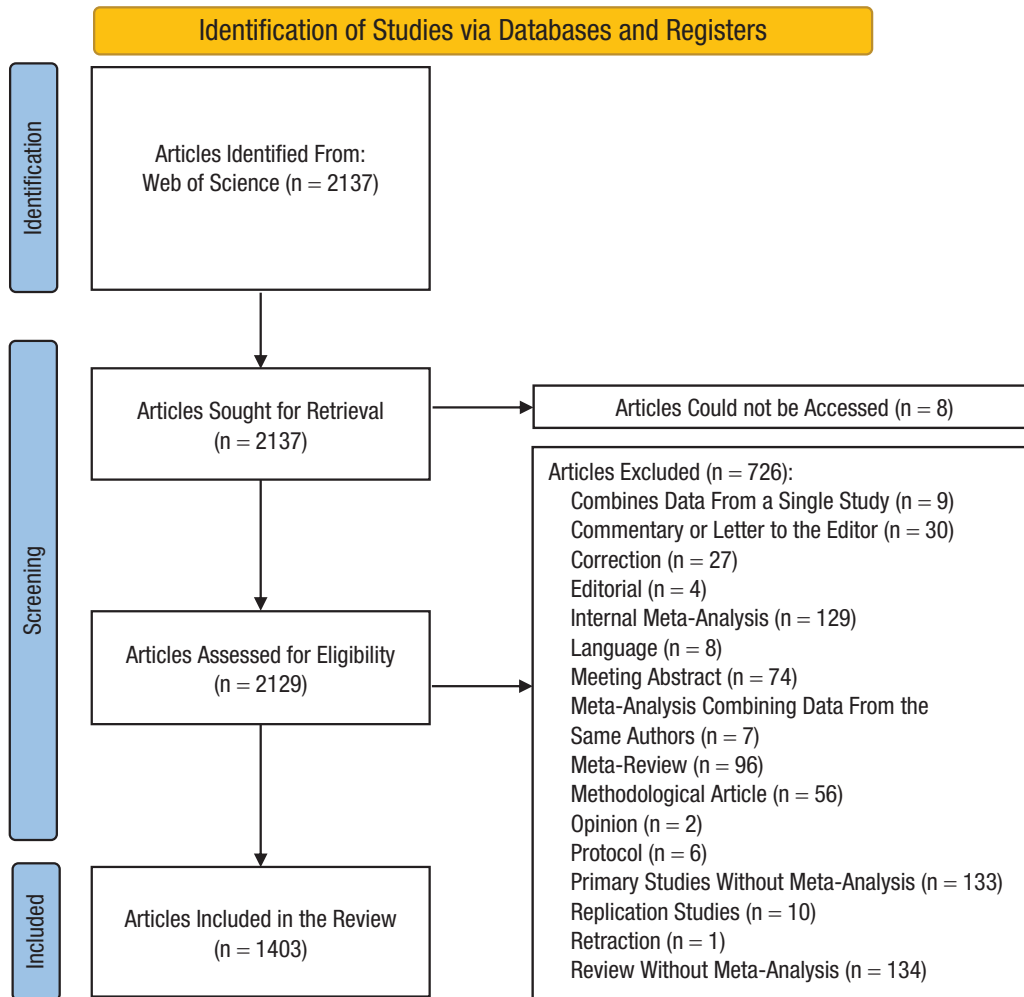
### Prevalence

Of the 1,403 eligible articles, 1,008 (72%) did not report having a preregistered protocol. Thirteen (1%) reported that they were preregistered, but we could not verify this because eight did not provide a link or other identifier, four provided a link that did not lead to any preregistered information or files, and one was registered after the article was submitted to a journal. Ultimately, we identified 382 (27%) articles that were associated with a functional preregistration.

The majority of the 382 preregistered articles were registered in the PROSPERO registry (308 articles; 81%), followed by the OSF registry (62 articles; 16%), INPLASY registry (five articles; 1.3%), both OSF and PROSPERO registries (three articles; 0.8%), both the OSF registry and PsyArXiv (one article; 0.3%), the journal *Systematic Reviews* (one article; 0.3%), Research Registry (one article; 0.3%), and the IRCT registry (one article; 0.3%).

### Coverage

Figure 2 shows the percentage of preregistered protocols that covered each PRISMA-P decision item. For a more detailed explanation of decision items, see Supplemental Table 1 at <https://osf.io/jq792>. In addition, supplemental figures (<https://osf.io/683z7>) provide a more detailed view of the coverage, deviations, and disclosure for each decision item using parallel-sets plots. Out of the 23 decisions items evaluated, preregistered protocols covered a median of 13 items (interquartile range [IQR] = 11–14). Information regarding the objectives (100%; 95% CI = [96%, 100%]), sources (100%; 95% CI = [96%, 100%]), eligibility criteria (99%; 95% CI = [95%, 100%]), and outcomes of interest (99%; 95% CI = [95%, 100%]) was present in almost every protocol. Likewise, risk of bias assessments (86%; 95% CI = [78%, 94%]), items to be extracted from primary studies (82%; 95% CI = [73%, 88%]), and additional analyses, such as moderator analyses or outlier analyses (80%; 95% CI = [71%, 87%]), were preregistered in most studies. Summary measures (72%; 95% CI = [63%, 80%]), screening process (70%; 95% CI = [60%, 78%]), searched dates (69%; 95% CI = [59%, 77%]), data-extraction process (65%; 95% CI = [55%, 73%]), and meta-analytic models (63%; 95% CI = [55%, 73%]) were also detailed in most studies. Information regarding dealing with missing data (51%; 95% CI = [41%, 61%]) and assessing heterogeneity (49%; 95% CI = [39%, 59%]) was present in about half of the protocols. Assessing publication bias (37%; 95% CI = [28%, 47%]), specific search strategies (37%; 95% CI = [28%, 47%]), and whether there was some kind of outcome prioritization (36%; 95% CI = [27%, 46%]) were not usually preregistered. Finally, methods for dealing with the possibility of multiple effect sizes from the same primary study (18%; 95%



**Fig. 1.** Preferred Reporting Items for Systematic review and Meta-Analysis 2020 flow diagram of the article selection process.

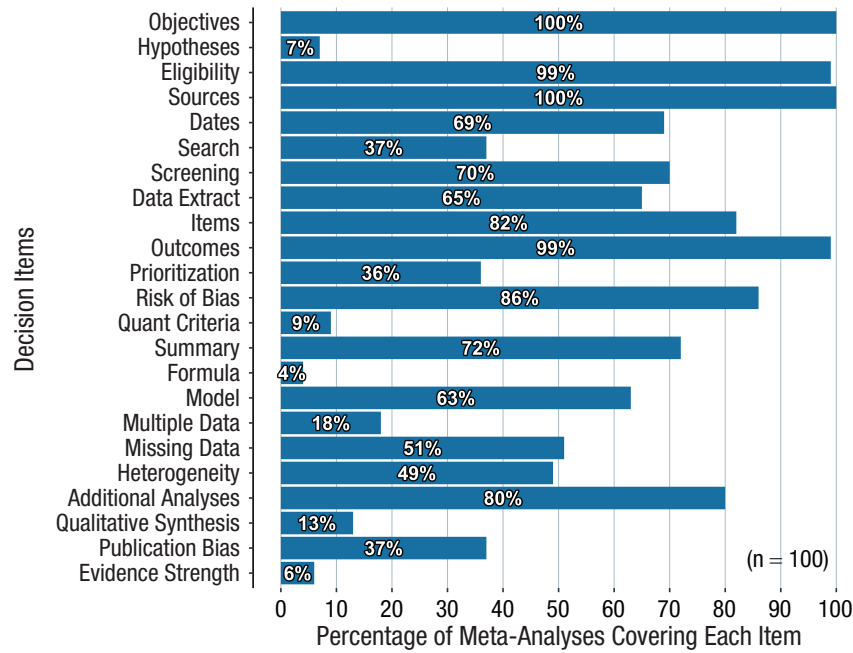
CI = [12%, 27%]), how a qualitative synthesis would be carried out if a quantitative analysis was not appropriate (13%; 95% CI = [27%, 46%]), the criteria for a primary study being included in a quantitative analysis (9%; 95% CI = [5%, 16%]), hypotheses (7%; 95% CI = [3%, 14%]), evaluating the strength of the evidence (6%; 95% CI = [3%, 12%]), and formula to compute effect sizes (4%; 95% CI = [2%, 10%]) were rarely preregistered.

**Deviations**

We found that every meta-analysis contained at least one deviation from the preregistered protocol, and there was a median of nine deviations (IQR = 6.75–11) per meta-analysis (see histogram in Supplemental Figure 1 at <https://osf.io/4wqzv>). Figure 3 shows the percentage of deviations for each decision item and the type of deviation. As noted above, supplemental figures (<https://osf.io/683z7>) provide a more detailed view of the coverage,

deviations, and disclosure for each decision item using parallel-sets plots. The highest rates of deviations were found regarding decisions items such as additional analyses (87%; 95% CI = [79%, 92%]), publication bias (66%; 95% CI = [56%, 74%]), items to be extracted from primary studies (60%; 95% CI = [50%, 69%]), assessment of heterogeneity (59%; 95% CI = [49%, 68%]), summary measures (58%; 95% CI = [48%, 67%]), search strategies (57%; 95% CI = [47%, 66%]), sources to retrieve the primary studies (56%; 95% CI = [46%, 65%]), and eligibility criteria (53%; 95% CI = [43%, 62%]).

In the case of additional analyses, eligibility criteria, sources, information to be extracted, and summary measures, these deviations were mainly because information was present in both protocol and final article but with modifications—that is, decision items were specified differently in the protocol and the final article (for an example of each type of deviation, see Table 1). Deviations regarding the assessment of publication bias,

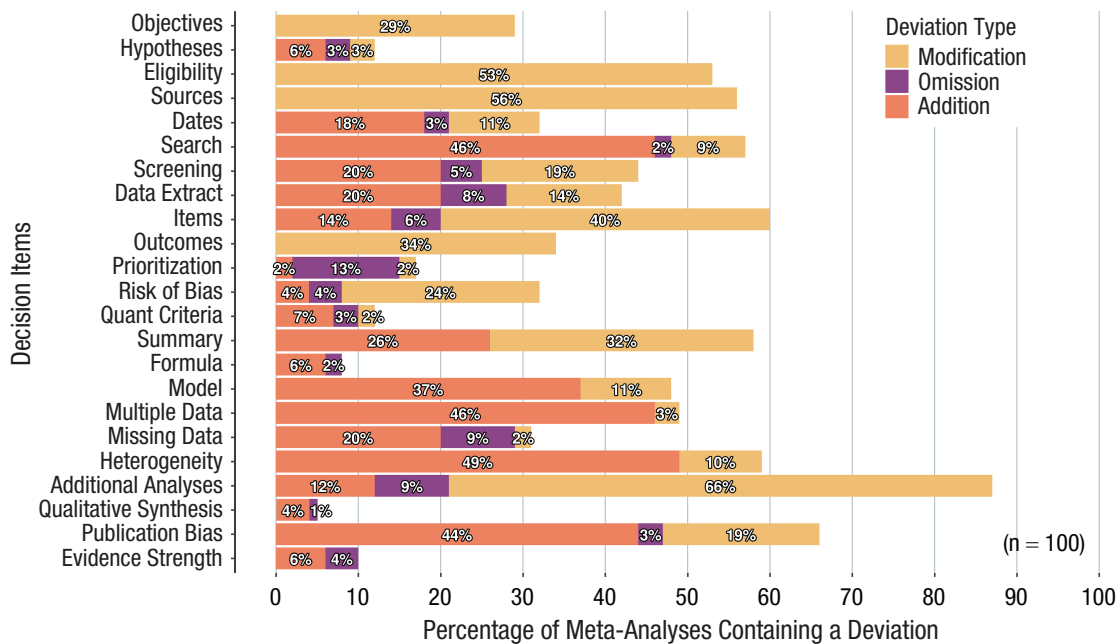


**Fig. 2.** Bar plot showing the percentage of preregistered meta-analysis protocols ( $N = 100$ ) that covered each decision item. Confidence intervals are reported in the main text.

heterogeneity, and search strategies were mainly due to information being present in the final article but not specified in advance in the protocol, that is, additions (see Table 1).

We also found frequent deviations for decision items related to dealing with multiple data from the same

primary study (49%; 95% CI = [39%, 59%]), statistical model (48%; 95% CI = [38%, 58%]), screening process (44%; 95% CI = [35%, 54%]), data-extraction process (42%; 95% CI = [33%, 52%]), outcomes of interest (34%; 95% CI = [25%, 44%]), dates that would be searched (32%; 95% CI = [24%, 42%]), assessment of the risk of



**Fig. 3.** Bar plot showing the percentage of meta-analyses ( $N = 100$ ) containing deviations for each decision item and the deviation type (modification, omission, addition). Confidence intervals are reported in the main text.

**Table 1.** Examples From Our Sample for the Different Types of Protocol Deviations (Modification, Omission, and Addition)

Type of deviation	Decision item	Protocol	Final article	Deviation information
Modification	Publication bias	Publication bias will also be assessed through investigation of funnel plots, trim and fill analyses and Egger's regression. (Source: Lorimer et al., 2019)	Potential publication bias was examined through investigation of funnel plots using Egger's regression and rank correlation tests. (Source: Lorimer et al., 2021)	The protocol and the final article both cover the decision item "publication bias"; however, the decision has been modified—trim and fill was not included in final article, and rank correlation tests were not included in the protocol.
Omission	Prioritization of outcomes	Main outcomes: BMI AN symptoms Additional outcomes: Anxiety Depression Psychosocial Adjustment (Source: <a href="https://osf.io/m9db2">https://osf.io/m9db2</a> )	Not mentioned (Source: Quiles Marcos et al., 2021)	The protocol differentiates between primary and secondary outcomes, but outcomes are not differentiated in the final article.
Addition	Risk of bias	Not mentioned (Source: Hou & Ye, 2020)	The approach developed by Hawker et al. (2002) was adopted to assess the methodological quality of individual studies included in this review. (Source: Ye et al., 2021)	The final article includes a risk of bias assessment, but the protocol does not.

Note: BMI = body mass index; AN = Anorexia Nervosa.

bias (32%; 95% CI = [24%, 42%]), dealing with missing data (31%; 95% CI = [23%, 41%]), and objectives (29%; 95% CI = [21%, 39%]).

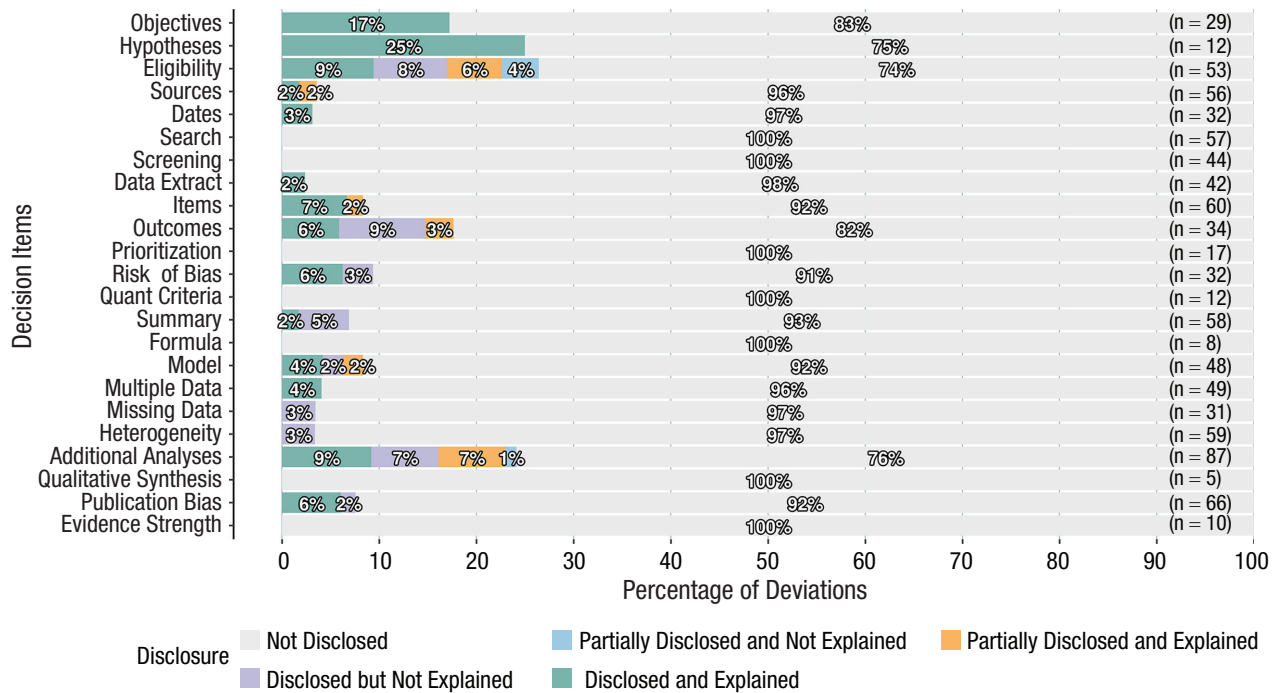
In the case of the statistical model, screening, data extraction, dates, and missing data, these deviations were mainly because information regarding these decision items was not present in the protocol but was later specified in the final article (additions). In turn, deviations regarding multiplicity of data, outcomes, and objectives were almost always in the form of modifications.

Finally, we also found deviations in the prioritization of outcomes (17%; 95% CI = [11%, 26%]), hypotheses (12%; 95% CI = [7%, 20%]), criteria for the studies to be included in the quantitative analysis (12%; 95% CI = [7%, 20%]), assessment of the strength of the evidence (10%; 95% CI = [6%, 17%]), and formulas to compute the effect sizes (8%; 95% CI = [4%, 15%]); description of the qualitative synthesis in a case study was not included in the quantitative synthesis (5%; 95% CI = [2%, 11%]).

In the case of the prioritization of outcomes, the deviations were mainly due to omissions because some kind of prioritization was specified in the protocol, but this

was not mentioned in the final article (see Table 1). For the hypotheses, criteria for quantitative analysis, assessment of the strength of the evidence, formulas, and description of a qualitative synthesis, deviations were mostly due to additions because protocols did not specify any information about these decision items but final articles did.

For most decision items, deviations were rarely disclosed. Every meta-analysis contained at least one undisclosed deviation, with a median of eight undisclosed deviations (IQR = 6–11) per meta-analysis (see histogram in Supplemental Figure 2 at <https://osf.io/4wqzv>). Figure 4 shows the percentage of deviations that were disclosed for each decision item. The deviations that were disclosed most frequently were those regarding the eligibility criteria (26%; 95% CI = [16%, 40%]), hypotheses (25%; 95% CI = [9%, 53%]), additional analyses (24%; 95% CI = [16%, 34%]), outcomes of interest (18%; 95% CI = [8%, 36%]), and objectives (17%; 95% CI = [8%, 35%]). When deviations in objectives and hypotheses were disclosed, authors always provided an explanation. However, when deviations in eligibility criteria, outcomes, or additional



**Fig. 4.** Bar plot showing the percentage of deviations for each decision item that were fully disclosed and explained, fully disclosed without any explanation, only partially disclosed but with explanation, partially disclosed without any explanation, and not disclosed ( $n$  indicates the number of deviations in each item). Confidence intervals are reported in the main text.

analyses were disclosed, authors sometimes disclosed only some but not all the deviations and/or did not always provide an explanation.

The rest of the decision-item deviations were all disclosed less than 10% of the time, and in some cases, the manuscript did not address all the deviations that we encountered or provide a justification for those deviations. Deviations regarding search strategies, screening, prioritization of outcomes, criteria for a quantitative synthesis, formulas to compute effect sizes, description of a qualitative synthesis, and assessment of the strength of the evidence were not discussed in any article.

## Discussion

In this study, we assessed the preregistration of meta-analyses in psychology. Of 1,403 eligible meta-analyses published in psychology in 2021, 382 (27%) were preregistered. This implies that the remaining 77% were exposed to a risk of bias from data-dependent decision-making. In a random sample of 100 preregistered meta-analyses, we found that key PRISMA-P decision items were often omitted from preregistered protocols—the median number of items covered was 13 (IQR = 11–14). We also found that all 100 assessed meta-analyses contained at least one deviation from the preregistered protocol ( $Mdn = 9$ ; IQR = 6.75–11). This implies that even preregistered meta-analyses incur a risk of bias from

data-dependent decision-making because they do not preregister key research decisions or indicate when they deviate from the preregistered plan. Concerningly, most deviations were undisclosed ( $Mdn = 8$ ; IQR = 6–11), which could potentially mislead readers and undermine their ability to calibrate their confidence in the reported results.

To our knowledge, these findings provide the first field-wide examination of the prevalence and practice of preregistration in psychology meta-analyses. Other studies have provided a range of estimates (2%–36%) for the prevalence of meta-analysis preregistration in various subfields of psychology (López-Nicolás et al., 2022; Polanin et al., 2020; Sandoval-Lentisco et al., 2024). Direct comparison is difficult because of differences in timeframe; however, considering these studies and our results, there is some evidence of a positive trend in meta-analysis preregistration.<sup>4</sup> Comparing the prevalence measure in the current study (27%) to a recent field-wide estimate (7%) for the prevalence of preregistration in empirical psychology (Hardwicke et al., in press) suggests that preregistration is more common for meta-analyses relative to empirical studies in general in psychology. This may be due to the influence of reporting guidelines such as PRISMA and organizations such as Cochrane (both of which recommend that meta-analyses be preregistered) in the community of psychologists who conduct meta-analyses and the existence

of a dedicated platform (PROSPERO) specifically focused on preregistrations of systematic reviews and meta-analyses (PROSPERO was the most popular registry among the preregistered meta-analyses we identified in this study).

Our findings of poor coverage of key research decisions in preregistered meta-analysis protocols and high frequency of undisclosed deviations align with other metaresearch studies examining preregistration of primary studies (Bakker et al., 2020; Claesen et al., 2021; TARG Meta-Research Group & Collaborators, 2023; van den Akker et al., 2023) and systematic reviews in medicine (Kirkham et al., 2010; Page et al., 2014; Siebert et al., 2023; Tricco et al., 2016). Taken together, these findings imply serious shortfalls in the implementation of preregistration that compromise its ability to minimize risk of bias and transparently communicate such risk to readers. Nonetheless, at the same time, these studies also illustrate the usefulness of preregistration as a tool that facilitates identification of risk of bias.

Considering the results shown here, we have some recommendations for researchers conducting meta-analyses in psychology. First, we encourage authors to be more specific and exhaustive in their protocols and pay special attention to decision items that we found to be frequently omitted from preregistered protocols (e.g., how to assess the presence of publication bias, deal with the multiplicity of effect sizes, assess the heterogeneity, and specify the meta-analysis model that will be assumed). To ensure the coverage of important decisions in preregistered protocols, authors can use tools such as PRISMA-P (Moher et al., 2015), which provides a checklist of the key decision items to be included in most meta-analyses. However, it is important to acknowledge that neither PRISMA-P nor PROSPERO provides complete coverage of all meta-analytic decisions, and therefore, authors need to specify more information than what these tools require by default. For better planning of the analyses, it has been suggested that authors simulate data to better understand the decisions to be made (Gambarota & Altoè, 2024; Nosek et al., 2019). Nonetheless, we acknowledge that specifying all meta-analytic decisions in advance can be difficult because many decisions depend on the characteristics of the data at hand (e.g., the choice of the meta-analysis model may depend on whether there are statistically dependent effect sizes). In this regard, handbooks (e.g., Cooper et al., 2019; Higgins et al., 2023) and guidelines (e.g., Carter et al., 2019; López-López et al., 2018; Nakagawa et al., 2022) may help to create prespecified decision rules for different scenarios that may occur.

Even if the preregistered protocol has been carefully specified, deviating from the plan may be necessary to maintain the validity of the analyses (Nosek et al., 2019). When deviations occur, authors should clearly disclose

them and provide a justification to help readers, reviewers, and editors understand why such decisions were made and the extent to which they incur a risk of bias. Some journals require preregistration deviations to be prominently disclosed and seek dedicated peer review of preregistrations (Hardwicke & Vazire, 2024; TARG Meta-Research Group & Collaborators, 2022). Authors might also consider using deviation-disclosure tables to clearly communicate deviations and their implications to readers (Moreau & Gamble, 2022; Willroth & Atherton, 2024). The Registered Report format might improve deviation disclosures (Claesen et al., 2021; Serghiou et al., 2023) given their emphasis on compliance with the Stage 1 (preregistered) protocol (Chambers & Tzavella, 2022; Hardwicke & Ioannidis, 2018). Because protocols are peer reviewed, Registered Reports may also improve protocol coverage and quality, potentially resulting in fewer deviations. Finally, *multiverse* analyses (Steege et al., 2016; Voracek et al., 2019) may increase transparency and reduce the risk of data-dependent decisions in meta-analyses. In this approach, instead of selecting only one analytical pathway, researchers perform many combinations of theoretically reasonable analysis specifications.

### ***Limitations and Future Directions***

An important limitation of our study is that we did not evaluate the impact of the decision-item coverage and deviations that we found. Although the decision items we evaluated theoretically incur a risk of bias, it is possible that in practice, they did not have a substantial impact on the results or conclusions of the meta-analyses. Future studies could evaluate whether these discrepancies could lead to substantial changes in the results by conducting different analyses with the different specifications reported in the protocols and the final articles.

Another limitation is that 62 articles had more than one protocol version available. In 35 of these cases, the protocols were updated after the study was completed or submitted to the journal. In such cases, we inspected only the protocol versions that were registered before completing or submitting the study. Twenty-seven articles had several protocol versions before the study was completed. For these articles, we selected the latest version available before the data analysis began. We decided to do this because it is possible that after commencing the study, researchers thought that a certain decision should be done differently before that decision was actually implemented. Nonetheless, it is also possible that authors updated their protocols after decisions were implemented, changing already completed aspects to match what they finally did. If this happened, we could have missed some deviations. In this sense, our criterion can be considered conservative. Despite this, the large

number of deviations found makes it unlikely that authors were frequently adjusting their protocols retrospectively to hide these deviations.

Finally, the interrater agreement regarding some decision items, such as the deviations for the objectives or the eligibility criteria, was only moderate. This result highlights that evaluating whether a deviation occurred can be challenging in some cases.

## Conclusions

Preregistration can be a useful way to minimize researcher biases and increase transparency, but its effectiveness relies on protocols covering key research decisions and protocol deviations being disclosed. We found that most meta-analyses published in psychology in 2021 were not preregistered, implying they are at risk of bias from data-dependent decision-making. When meta-analyses were preregistered, their protocols often provided poor coverage of key research decisions, and researchers frequently deviated from them, often without disclosing that a deviation had occurred. Thus, even preregistered meta-analyses often incurred a risk of bias. In sum, although preregistration is useful to detect risk of bias, its infrequent use and poor implementation in psychology meta-analyses undermine its potential to reduce bias and increase transparency.

## Transparency

*Action Editor:* Pamela-Davis Kean

*Editor:* David A. Sbarra

*Author Contributions*

**Alejandro Sandoval-Lentisco:** Conceptualization; Formal analysis; Investigation; Methodology; Software; Writing – original draft; Writing – review & editing.

**Miriam Tortajada:** Investigation; Writing – review & editing.

**Rubén López-Nicolás:** Investigation; Writing – review & editing.

**José A. López-López:** Conceptualization; Funding acquisition; Writing – review & editing.

**E. J. Wagenmakers:** Conceptualization; Writing – review & editing.

**Julio Sánchez-Meca:** Conceptualization; Funding acquisition; Writing – review & editing.

**Tom E. Hardwick:** Conceptualization; Formal analysis; Methodology; Writing – original draft; Writing – review & editing.

*Declaration of Conflicting Interests*

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

*Funding*

This study was funded by Agencia Estatal de Investigación (Government of Spain), Fondo Europeo de Desarrollo Regional Funds (Project PID2019-104080GB-I00/MCIN/AEI/10.13039/501100011033), and the Region of Murcia

(Spain) through the Regional Program for the Promotion of Scientific and Technical Research of Excellence (Action Plan 2022) of the Seneca Foundation - Science and Technology Agency of the Region of Murcia (Grant 22064/PI/22).

*Open Practices*


This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.




## ORCID iDs

Alejandro Sandoval-Lentisco  <https://orcid.org/0000-0002-7876-0101>

Rubén López-Nicolás  <https://orcid.org/0000-0002-6963-7443>

E. J. Wagenmakers  <https://orcid.org/0000-0003-1596-1034>

Julio Sánchez-Meca  <https://orcid.org/0000-0002-8412-788X>

Tom E. Hardwick  <https://orcid.org/0000-0001-9485-4952>

## Acknowledgment

We thank Robert Thibault for providing valuable feedback at an early stage of the study.

## Notes

- Note that we did not require meta-analyses to have followed the PRISMA-P guidelines. We simply used this tool as a guide to which aspects we should assess.
- PRISMA-P did not include any item related to hypotheses; however, we included hypotheses as a decision item because added/changed/omitted hypotheses can create a risk of bias.
- Note that here we refer to biases introduced by the researchers conducting a meta-analysis (Whiting et al., 2016) and not the risk of bias in primary studies included in the meta-analysis.
- Sandoval-Lentisco et al. (2024) reported a positive relationship between publication year and preregistration, considering both the whole sample of meta-analysis but also selecting only those meta-analyses published since 2011, the year in which the PROSPERO platform was launched. Furthermore, we conducted another analysis using a binary logistic regression between year and preregistration using the data from López-Nicolás et al. (2022). We also found a positive trend considering both every meta-analysis published from 2000 and meta-analyses published after 2011.

## References

- Bakker, M., Veldkamp, C. L. S., van Assen, M. A. L. M., Crompvoets, E. A. V., Ong, H. H., Nosek, B. A., Soderberg, C. K., Mellor, D., & Wicherts, J. M. (2020). Ensuring the quality and specificity of preregistrations. *PLOS Biology*, *18*(12), Article e3000937. <https://doi.org/10.1371/journal.pbio.3000937>
- Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and*

- Practices in Psychological Science*, 2(2), 115–144. <https://doi.org/10.1177/2515245919847196>
- Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. *Nature Human Behaviour*, 6(1), Article 1. <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(10), Article 211037. <https://doi.org/10.1098/rsos.211037>
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2019). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation.
- Deeks, J. J., Higgins, J. P., & Altman, D. G. (2023). Analysing data and undertaking meta-analyses. In J. P. T. Higgins & J. Thomas (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 243–296). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470712184.ch9>
- Dwan, K., Kirkham, J. J., Williamson, P. R., & Gamble, C. (2013). Selective reporting of outcomes in randomised controlled trials in systematic reviews of cystic fibrosis. *BMJ Open*, 3(6), Article e002709. <https://doi.org/10.1136/bmjopen-2013-002709>
- Gambarota, F., & Altoè, G. (2024). Understanding meta-analysis through data simulation with applications to power analysis. *Advances in Methods and Practices in Psychological Science*, 7(1). <https://doi.org/10.1177/25152459231209330>
- Hardwicke, T. E., & Ioannidis, J. P. A. (2018). Mapping the universe of Registered Reports. *Nature Human Behaviour*, 2(11), 793–796. <https://doi.org/10.1038/s41562-018-0444-y>
- Hardwicke, T. E., Thibault, R. T., Clarke, B., Moodie, N., Crüwell, S., Schiavone, S. R., Handcock, S., Mody, F., Europe, T., & Vazire, S. (in press). Prevalence of transparent research practices in empirical psychology articles published in 2022: A cross sectional study. *Advances in Methods and Practices in Psychological Science*.
- Hardwicke, T. E., Thibault, R. T., Kosie, J. E., Wallach, J. D., Kidwell, M. C., & Ioannidis, J. P. (2022). Estimating the prevalence of transparency and reproducibility-related research practices in psychology (2014–2017). *Perspectives on Psychological Science*, 17(1), 239–251.
- Hardwicke, T. E., & Vazire, S. (2024). Transparency is now the default at psychological science. *Psychological Science*, 35(7), 703–824. <https://doi.org/10.1177/09567976231221573>
- Hardwicke, T. E., & Wagenmakers, E.-J. (2023). Reducing bias, increasing transparency and calibrating confidence with preregistration. *Nature Human Behaviour*, 7(1), Article 1. <https://doi.org/10.1038/s41562-022-01497-2>
- Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., & Ioannidis, J. P. (2020). An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017). *Royal Society Open Science*, 7(2), 190806.
- Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2023). *Cochrane handbook for systematic reviews of interventions version 6.4 (updated August 2023)*. Cochrane.
- Hou, R., & Ye, G. (2020). *Anxiety in asthma: A systematic review and meta-analysis*. PROSPERO 2020 CRD42020176028. [https://www.crd.york.ac.uk/prospERO/display\\_record.php?ID=CRD42020176028](https://www.crd.york.ac.uk/prospERO/display_record.php?ID=CRD42020176028)
- Johnson, B. T. (2021). Toward a more transparent, rigorous, and generative psychology. *Psychological Bulletin*, 147(1), 1–15. <https://doi.org/10.1037/bul0000317>
- Kirkham, J. J., Altman, D. G., & Williamson, P. R. (2010). Bias due to changes in specified outcomes during the systematic review process. *PLOS ONE*, 5(3), Article e9810. <https://doi.org/10.1371/journal.pone.0009810>
- López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. T. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods*, 9(3), 336–351. <https://doi.org/10.1002/jrsm.1310>
- López-Nicolás, R., López-López, J. A., Rubio-Aparicio, M., & Sánchez-Meca, J. (2022). A meta-review of transparency and reproducibility-related reporting practices in published meta-analyses on clinical psychological interventions (2000–2020). *Behavior Research Methods*, 54(1), 334–349. <https://doi.org/10.3758/s13428-021-01644-z>
- Lorimer, B., Delgadillo, J., Kellett, S., & Nye, A. (2019). *Predictors of relapse after acute-phase cognitive behavioural therapy for anxiety disorders, post-traumatic stress disorder, and obsessive-compulsive disorder: A systematic review*. PROSPERO 2019 CRD42019133033. [https://www.crd.york.ac.uk/prospERO/display\\_record.php?ID=CRD42019133033](https://www.crd.york.ac.uk/prospERO/display_record.php?ID=CRD42019133033)
- Lorimer, B., Kellett, S., Nye, A., & Delgadillo, J. (2021). Predictors of relapse and recurrence following cognitive behavioural therapy for anxiety-related disorders: A systematic review. *Cognitive Behaviour Therapy*, 50(1), 1–18.
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A., & PRISMA-P Group. (2015). Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, Article 1. <https://doi.org/10.1186/2046-4053-4-1>
- Moreau, D., & Gamble, B. (2022). Conducting a meta-analysis in the age of open science: Tools, tips, and practical recommendations. *Psychological Methods*, 27(3), 426–432. <https://doi.org/10.1037/met0000351>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), Article 1. <https://doi.org/10.1038/s41562-016-0021>
- Nakagawa, S., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W. A., Parker, T. H., Sánchez-Tójar, A., Yang, Y., & O’Dea, R. E. (2022). Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*, 13(1), 4–21. <https://doi.org/10.1111/2041-210X.13724>
- Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., van ’t Veer, A. E., & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, 23(10), 815–818. <https://doi.org/10.1016/j.tics.2019.07.009>
- Page, M. J., McKenzie, J. E., Kirkham, J., Dwan, K., Kramer, S., Green, S., & Forbes, A. (2014). Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions.

- Cochrane Database of Systematic Reviews*, 10. <https://doi.org/10.1002/14651858.MR000035.pub2>
- Parmelli, E., Liberati, A., & D'Amico, R. (2007). *Reporting of outcomes in systematic reviews: Comparison of protocols and published systematic reviews*. Colloquium Abstracts. <https://abstracts.cochrane.org/2007-sao-paulo/reporting-outcomes-systematic-reviews-comparison-protocols-and-published-systematic>
- Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, 90(1), 24–46. <https://doi.org/10.3102/0034654319877153>
- Polanin, J. R., Hennessy, E. A., & Tsuji, S. (2020). Transparency and reproducibility of meta-analyses in psychology: A meta-review. *Perspectives on Psychological Science*, 15(4), 1026–1041. <https://doi.org/10.1177/1745691620906416>
- Quiles Marcos, Y., Leon Zarceno, E., & Lopez Lopez, J. A. (2021). Effectiveness of exercise-based interventions in patients with anorexia nervosa: A systematic review. *European Eating Disorders Review*, 29(1), 3–19.
- Sandoval-Lentisco, A., López-Nicolás, R., Tortajada, M., López-López, J. A., & Sánchez-Meca, J. (2024). Transparency in cognitive training meta-analyses: A meta-review. *Neuropsychology Review*. Advance online publication. <https://doi.org/10.1007/s11065-024-09638-2>
- Serghiou, S., Axfors, C., & Ioannidis, J. P. A. (2023). Lessons learnt from registration of biomedical research. *Nature Human Behaviour*, 7(1), Article 1. <https://doi.org/10.1038/s41562-022-01499-0>
- Siebert, M., Caquelin, L., Madera, M., Acosta-Dighero, R., Naudet, F., & Roqué, M. (2023). Assessing the magnitude of changes from protocol to publication—A survey on Cochrane and non-Cochrane Systematic Reviews. *PeerJ*, 11, Article e16016. <https://doi.org/10.7717/peerj.16016>
- Silagy, C. A., Middleton, P., & Hopewell, S. (2002). Publishing protocols of systematic reviews: Comparing what was done to what was planned. *JAMA*, 287(21), 2831–2834. <https://doi.org/10.1001/jama.287.21.2831>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Sison, C. P., & Glaz, J. (1995). Simultaneous confidence intervals and sample size determination for multinomial proportions. *Journal of the American Statistical Association*, 90(429), 366–369. <https://doi.org/10.1080/01621459.1995.10476521>
- Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, 11(5), 702–712. <https://doi.org/10.1177/1745691616658637>
- TARG Meta-Research Group & Collaborators. (2022). Discrepancy review: A feasibility study of a novel peer review intervention to reduce undisclosed discrepancies between registrations and publications. *Royal Society Open Science*, 9(7), Article 220142. <https://doi.org/10.1098/rsos.220142>
- TARG Meta-Research Group & Collaborators. (2023). Estimating the prevalence of discrepancies between study registrations and publications: A systematic review and meta-analyses. *BMJ Open*, 13(10), Article e076264. <https://doi.org/10.1136/bmjopen-2023-076264>
- Tricco, A. C., Cogo, E., Page, M. J., Polisen, J., Booth, A., Dwan, K., MacDonald, H., Clifford, T. J., Stewart, L. A., Straus, S. E., & Moher, D. (2016). A third of systematic reviews changed or did not specify the primary outcome: A PROSPERO register study. *Journal of Clinical Epidemiology*, 79, 46–54. <https://doi.org/10.1016/j.jclinepi.2016.03.025>
- van den Akker, O. R., van Assen, M. A. L. M., Enting, M., de Jonge, M., Ong, H. H., Rüffer, F., Schoenmakers, M., Stoevenbelt, A. H., Wicherts, J. M., & Bakker, M. (2023). Selective hypothesis reporting in psychology: Comparing preregistrations and corresponding publications. *Advances in Methods and Practices in Psychological Science*, 6(3). <https://doi.org/10.1177/25152459231187988>
- Voracek, M., Kossmeier, M., & Tran, U. S. (2019). Which data to meta-analyze, and how? *Zeitschrift Für Psychologie*, 227(1), 64–82. <https://doi.org/10.1027/2151-2604/a000357>
- Whiting, P., Savović, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., Churchill, R., & ROBIS group. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69, 225–234.
- Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology*, 7, Article 1832. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01832>
- Willroth, E. C., & Atherton, O. E. (2024). Best laid plans: A guide to reporting preregistration deviations. *Advances in Methods and Practices in Psychological Science*, 7(1). <https://doi.org/10.1177/25152459231213802>
- Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 209–212. <https://doi.org/10.1080/01621459.1927.10502953>
- Ye, G., Baldwin, D. S., & Hou, R. (2021). Anxiety in asthma: A systematic review and meta-analysis. *Psychological Medicine*, 51(1), 11–20.