

Guidelines for conducting and publishing systematic reviews in Psychology

Diretrizes para condução e publicação de revisões sistemáticas em Psicologia

Lucas de Francisco **CARVALHO**¹  0000-0002-3274-9724

Giselle **PIANOWSKI**¹  0000-0003-4400-7151

Manoel Antônio dos **SANTOS**²  0000-0001-8214-7767

Abstract

The systematic review refers to the literature review guided by scientific methods explicitly intended to reduce bias, resulting in a synthesis of all relevant evidence for a given issue. In Brazil, specifically in Psychology, systematic review is found in the literature; however, the available studies do not always reflect the gold standard or what is expected in terms of typical systematic review procedures. The present study is structured in the form of a didactic guide, organized in topics, which should be typically contemplated in an systematic review in Psychology. The information that must be contained in each of these topics is indicated, including which procedures should be performed in the typical steps of the development of an systematic review. The present publication intends to increase the interest and investment of researchers in systematic review, providing them with information to improve the quality of systematic review in the area of Psychology in Brazil.

Keywords: Investigation; Meta-analysis; Research design; Review literature.

Resumo

A revisão sistemática refere-se à revisão de literatura guiada por métodos científicos explicitamente voltados para redução de vieses, resultando em uma síntese de toda evidência relevante para uma área. No Brasil, especificamente na Psicologia, revisões sistemáticas são encontradas na literatura; entretanto, os estudos disponíveis nem sempre refletem

▼ ▼ ▼ ▼ ▼

¹ Universidade São Francisco, Programa de Pós-Graduação em Psicologia. R. Waldemar César da Silveira, 105, Jardim Cura D'Ars, 13045-510, Campinas, SP, Brasil. *Correspondência para/Correspondence to:* L.F. CARVALHO. E-mail: <lucas@labape.com.br>.

² Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Departamento de Psicologia. Ribeirão Preto, SP, Brasil.

▼ ▼ ▼ ▼ ▼

Como citar este artigo/How to cite this article

Carvalho, L. F., Pianowski, G., & Santos, M. A. (2019). Guidelines for conducting and publishing systematic reviews in Psychology. *Estudos de Psicologia* (Campinas), 36, e180144. <http://dx.doi.org/10.1590/1982-0275201936e180144>



o padrão ouro ou o que é esperado em relação a tais procedimentos. Nesse sentido, este estudo é estruturado na forma de guia didático e organizado em tópicos, apresentando o que deve ser contemplado em revisões sistemáticas na área de Psicologia. A informação que deve ser contida em cada tópico é apresentada, incluindo notas breves acerca de quais procedimentos devem ser realizados em uma revisão sistemática. A presente publicação intenciona aumentar o interesse e investimento de pesquisadores em relação à matéria, fornecendo informações para melhoria da qualidade das revisões sistemáticas na área da Psicologia no Brasil.

Palavras-chave: *Investigação; Metanálise; Projetos de pesquisa; Revisão de literatura.*

Publications that use the method of systematic review of the literature, or only Systematic Review (SR) as it is better known, are typically found in the scientific literature, covering the most diverse areas of knowledge. The use of SR is linked to decision-making (Higgins & Green, 2011; Petticrew, 2001) for different scientific issues, including those with greater direct impact on society, public policies and professional practice, as well as those with an indirect impact, more related to basic research, laboratory studies, tool development, among other applications.

The literature (Holly, Salmond, & Saimbert, 2012; Littell, Corcoran, & Pillai, 2008; Petticrew & Roberts, 2006), points out that the first SR was conducted by Nichols (1981) in the field of Psychology (American Journal of Psychology), still without the designation of the term (systematic review). It is, therefore, a relatively recent procedure when viewed from a historical perspective.

Since then, reviews have been made, many of them in the context of health sciences (Higgins & Green, 2011). An SR is understood as the systematic appraisal of a comprehensive body of research, related to subjects of interest for the development of a certain field of scientific knowledge. It is noteworthy that choosing not to perform an SR, in favor of other types of non-systematized revisions, can culminate in biases with practical implications. For example, Pauling (1987) published a book reporting that the use of megadoses of vitamin C would prevent the cold; years later, it was observed in an SR that megadoses of vitamin C tend not to be effective for cold prevention (Knipschild, 1994). In contrast to this unfavorable outcome, we can mention the example of a study that used an SR and that had direct implications in society and in the formulation of public policies (Treasure et al., 2002), which investigated the available evidence to support the decision on whether or not fluoride should be included in running water in England.

Several myths and misunderstandings surround the SR, so researchers interested in this procedure should seek reliable sources for proper understanding. An SR can be defined as a literature review guided by scientific methods explicitly intended to reduce bias, resulting in a synthesis of all relevant evidence for a particular issue (Higgins & Green, 2011; Holly et al., 2012; Petticrew & Roberts, 2006). Some considerations should be highlighted: SR refers to a scientific tool for literature reviews; the results of an SR should guide the next steps of a given area; the SR should present a conclusion, resulting in the delimitation of the state of the art of the knowledge produced; the SR starts from the presupposition of evidence-based practice; the result obtained on the basis of an SR is less susceptible to bias in comparison to primary studies.

Some of the main benefits that an SR offers are: to address in detail the specificities of a given problem; to aggregate a large amount of information; to achieve high scientific rigor in dealing with the evidence; to reduce the impact of biases of systematic and non-systematic errors; to synthesize relevant information for different areas; to focus on the consistencies of the studies, pondering their inconsistencies; to increase statistical power over primary studies; to engender more “credibility” on the conclusion reached; to enable the updating and evolution of the area in which the research problem is inserted; and to facilitate updating by professionals in the field investigated (Higgins & Green, 2011; Holly et al., 2012).

An extensive literature is available on the definitions, concepts and applications of the SR (Higgins & Green, 2011), as well as studies that discuss the most frequent myths surrounding it (Petticrew, 2011). In the area of health, primarily in the medical and pharmaceutical fields, Cochrane plays a relevant role in the

development of guidelines and improvement of methods for conducting SRs and meta-analyses, besides providing a repository for performed reviews and relevant technical material to guide the work of review authors. The intervention area, in this case, is given priority (Higgins & Green, 2011). Practical guides and guidelines, which give a comprehensive description of how to conduct an SR and how to disseminate it, are scarcer for some other areas of science. Specifically, Cochrane is a British charitable organization created to organize medical research findings to facilitate evidence-based decisions. The organization conducts SRs focused on health-care interventions and diagnostic tests, and publishes them in the Cochrane Library. According to Grimshaw (2004), Cochrane reviews are more up to date, with better quality in comparison to other reviews, and are probably responsible for a significant portion of the methodological improvements in the medical literature.

In Brazil, specifically in Psychology, SRs are found in the literature; however, available studies do not always reflect the gold standard and/or what is expected in terms of typical SR procedures. Empirical evidence of this situation has been published in the literature (Zoltowski, Costa, Teixeira, & Koller, 2014), complaining about the poor quality of SRs performed in the national context in the area of psychology. Although large impact initiatives in the medical field are available (e.g., Cochrane), we could not find publications presenting specific guidelines for the area of psychology on how to conduct SRs. Given this setting, the present study aims to present and discuss the main guidelines for conducting and disseminating an SR in psychology. Extrapolating this scope, some specifics will also be presented for some SR cases followed by meta-analyses.

The present study is structured in the form of didactic guidelines, organized into topics, so that the topics that should be typically included in an SR in psychology are presented (Introduction, Method, Results, Discussion and Final Considerations). The paragraphs included in each of the topics describe what kind of information should be contained therein, and briefly, which procedures should be performed in each step, when applicable.

We highlight that the guidelines and references presented followed the prominent literature in the area, primarily the Cochrane Handbook (Higgins & Green, 2011), and the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement (McInnes et al., 2018), as well as the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) (Collins, Reitsma, Altman, & Moons, 2015). As no specific guidelines were found for SRs in psychology, we selected the information identified as pertinent to the area and object of study. We present below the first topic (i.e., Introduction) and the due recommendations.

Guidelines for systematic review

Introduction

Similarly to what is expected in scientific publications in psychology in general, in this topic the authors must justify the area of study in which the SR fits, as well as justify the realization of the proposal. If there are already SRs published in the specific area and topic, they can (and, in most cases, must) be presented, allowing the authors to position themselves and to provide the justifications (such as controlling some variable that was not controlled in the previous review) for a new SR.

The introduction should provide the reader with clear reasons that support the authors to detect the need to perform an SR with a given focus. For this purpose, the authors should consider that an SR is a tool that allows identification of the level of maturity of a research area; assists in tracing or mapping gaps in the literature; allows aggregating information to respond with less bias to a given research problem; provides support in decision-making (e.g., in professional and political activities); helps to propose future steps for

research and development of the area, among other purposes. Broadly understanding the scope of the SR can facilitate the development of the argument that supports and justifies the use of this procedure in a study.

The objective can be presented with a central focus on the systematic review (e.g., “the objective of this study was to perform an SR...”) or it can be presented with a focus on the research problem itself. In this case, the authors should explain that the SR was used as a proceeding (e.g., “the objective of this study was to investigate the relationship between personality traits and fluid intelligence, for this purpose we used an SR...”). In cases of SR with meta-analysis, authors should explicitly state this purpose in the objective.

The authors are expected to clearly define a guiding question for the review. Especially in cases of intervention-focused research, it is advisable to use the formats “Population, Intervention, Comparison and Outcomes” (PICO), or PICOT (adding Time) or PICOS (adding Study design) to prepare the guiding question. The question also guides the researcher in the analysis of scientific production, in which he/she must identify relations, contradictions, gaps and/or inconsistencies in the literature. (From the results, the researcher should also suggest the next steps to be pursued by future research to solve the problems identified).

Method

There are sites, platforms and software tools available to assist reviewers in building and performing a systematic review. However, most of them are intended for health science with experimental or quasi-experimental research methodology, besides a large number of directly measurable interventions and outcomes (e.g. Randomized Controlled Trials [RCTs]). Researchers who focus on psychological constructs in an observational/cross-sectional design (i.e., much of the research in psychology and mental health in general) should adapt such tools to the specific purpose. An example is Cochrane’s Confidence platform for literature management, data extraction, protocol creation, and peer review, among other functions.

It is necessary to clarify the criteria used in the selection of the studies to be included in the review, as well as the procedures used in the synthesis of the results obtained by the reviewed articles (which may or may not be the meta-analysis). The author of an SR will have to make careful decisions for a correct execution. When defining and describing the method used for an SR, care must be taken to focus on two central issues: rigor and replicability. As the SR is rigorously aimed at eliciting information about existing evidence on a particular issue (e.g., whether or not there is evidence supporting the use of the Personality Inventory for DSM-5 to assess antisocial personality disorder). Therefore, its method must be built and guided with the necessary rigor to ensure that results are comprehensive, yet specific and reliable.

In constructing and applying the proposed methodology, the author of the manuscript that reports the SR should also ensure that its readers can achieve the same results if they follow the reported methodological decisions. Next, we present some recommendations related to rigor and replicability in SR.

Initially, it is recommended that SRs be registered in the International Prospective Register of Systematic Reviews (PROSPERO), and that their registration be duly reported in the Method. This report can be inserted before the subtopics of the Method. PROSPERO is an international database for SR registration in several areas, including the health area in a broad way. It provides a list for public consultation of the SR performed, in progress or to be initiated, making it possible to compare what was proposed in the SR registry and what was actually done. In addition, the procedures to be followed, as well as the information that should be minimally presented, be established from recognized guidelines. In fact, a systematic review should describe in detail the methodology used to search for the original studies.

Among the main recommended guidelines, there is PRISMA (Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009; Moher et al., 2015), which concerns a group of information, in the form of items, to

be presented in an SR report (e.g., scientific article) or the Cochrane Collaboration. Following the PRISMA recommendations implies giving robustness to the SR.

After these steps, information on methodological procedures and decisions should be detailed and, preferably, organized into subtopics. This will be important to ensure accuracy during the execution of the SR, and to allow other researchers to replicate it. There is no specific format for this subdivision, but information to be presented includes: definition of survey strategy, eligibility criteria, extraction and management of data, evaluation procedures and selection of primary studies, quality assessment of the report and the risk of bias, assessment of publication bias, and statistical analyses. These last two topics are usually presented when the SR allows the performance of a meta-analysis. The topics mentioned are now presented.

Search strategy

Authors should indicate where and how the studies were surveyed. It is important to keep in mind that the search must be as sensitive as possible to collect the largest amount of accessible data, aiming to keep the descriptors specific enough not to cover too much irrelevant literature. Language(s) in which the survey was performed, the time period considered for the search, the date of the survey (it is currently recommended that searches be performed in the different indexing databases in a single day), which databases were used, how authors handled unpublished material and whether or not gray literature was sought (i.e., research results that were not published in the usual scientific publications and/or commercial channels) and which uniterms/descriptors were used in the survey, should be presented in this topic.

It is suggested that all SRs be performed primarily in the English language and that the author does not add a language restriction when the full content is not familiar (material in a language not familiar to the author should be translated). In contrast, there are rare cases that justify SR restricted to a language (other than English), e.g., when there is some cultural and/or contextual reason that clearly indicates that the pertinent literature is concentrated preferentially in a particular language. However, in most cases, the English language should be the key language for the execution of the SR, as English is the language adopted by science produced in a globalized world.

The search period, in turn, may be established from some key date (e.g., year of release of an instrument), or else insert no time constraint. What is important is that the date inserted should not restrict the scope of the search and should be justifiable. It is also necessary to register and inform when the survey was performed. It is not necessary that this information be detailed in day, month and year, being sufficient the indication of month and year.

Regarding the choice of databases and repositories for literature review, authors should decide whether to focus only on content already published in scientific channels (e.g., databases), whether they will cover unpublished and ongoing material (e.g., partial data published in conference annals) and whether the gray literature will be used (e.g., reports of institutions not published by scientific channels). It is recommended to give priority to the first channel (i.e., already published scientific material), but the concern in reaching material not published yet and gray literature may contribute to add value and increase the comprehensiveness of the review.

Data published in scientific channels should be retrieved on an index basis (a more detailed description will be given below). Unpublished or partially published materials (e.g., in progress) can be identified in preliminary data, such as in conference annals and others, and it is then possible to access the authors to request the complete or updated material (if any). The gray literature, in turn, refers to any report or content that is the result of rigorous research, but which has not been disseminated in channels intended specifically for scientific publication, such as reports by the government, world organizations, private companies, and

others. There are databases that are focused on surveying the gray literature, such as GreyNet International and OpenGrey. In health, especially in medicine, there is a significant amount of gray literature, which makes this survey relevant for many revisions. Psychology authors must rely on the expertise in the area of study to decide on whether to verify the existence and pertinence of incorporating the gray literature into the SR.

In relation to the databases, there is no gold standard on how many or which should be used. The determination of the bases depends essentially on the area in which the SR is found, but researchers must ensure that the main bases are being covered (e.g., PsycInfo, Scopus, PubMed, Web of Science). Furthermore, the use of Google Scholar has been increasingly recommended as a complementary basis to the main ones used (Gehanno, Rollin, & Darmoni, 2013; Haddaway, Collins, Coughlin, & Kirk, 2015; De Winter, Zadpoor, & Dodou, 2014). Where relevant, thesis and dissertation bases should be included, and care should be taken to check whether the content (partial or total) has been published in scientific channels (journals) and, if so, the reference found should be replaced by the corresponding article. It is understood that publications in scientific channels, *a priori*, should involve greater control of bias and of the quality of the report, which can contribute to ensure the adequacy of the results. Another option, which is often relevant for authors of SRs, is the use of hand-searching (e.g., searching for reference authors in the area; searching for sites that are not strictly scientific). Hand-searching may assist in the survey of gray literature and occasionally, scientific literature, and should be reported in a clear, coherent and replicable manner.

After defining and presenting the bases that were used, care is necessary to select the main descriptors and variations to insert in the search, informing them in the SR report. Whenever possible, it is recommended that researchers use the available mechanisms that assist in choosing the most commonly used terms. For example, for SRs in health, PubMed provides the Medical Subject Headings (MeSH), which allows, among other functions, to check which possible terms are available in different health research niches (e.g., for Borderline Personality Disorder, the MeSH suggests entering the following terms: Disorder, Borderline Personality, Borderline Personality Disorders, Disorders, Borderline Personality, Personality Disorders, Borderline, Personality Disorder, Borderline). Other systems that allow the identification of descriptors frequently used in the health and life areas are the Health Sciences Descriptors (HSD), the Medical Terminologies at National Library of Medicine (NLM), and the Elsevier's Life Science Thesaurus (EMTREE/EMBASE). The latter contains paid packages destined to the use of individual researchers and institutions. Many of the terms relevant to psychiatry are also relevant to psychology, which allows psychologist researchers to make a productive use of these tools. Furthermore, in psychology the Thesaurus is commonly used to define descriptors.

In addition, some databases (e.g., PubMed) aid in the entire process of constructing the search strategy (see PubMed Search Builder). This assistance is necessary so that the user can gain knowledge on the use of connectors and specifiers (e.g., AND, OR, NOT and ""). We emphasize the relevance of the correct use of connectors, abbreviations, symbols and specifiers according to particularities of each database (e.g., the use of * in some bases allows all derivations of a word to be considered, for example, *disorder**, for *disorder* and *disorders*).

Regarding the terms in the publication, some scientific journals allow the full publication of the search routines, even if extensive, either in the manuscript file itself or in supplementary material. When this is not the case, authors can choose to present only the main terms and report which variations have also been used (according to MeSH, for example). Considering that the contact of the researchers is informed in the publication, readers who want to access the search routines in full may establish direct contact with the authors.

Eligibility criteria

Eligibility criteria must be developed simultaneously to the goal and search strategy, because they can impact decisions of both descriptors and search sites. Reporting the eligibility criteria can be done soon after

the detailed description of the search strategy and definition of the publication bases used. In this topic the researchers should carefully present the inclusion and exclusion criteria that were used to select the publications. These criteria must be justified when they are atypical, that is, when there is a specific reason for adopting a given criterion, and that is not sufficiently clear in the SR context. For example, it is expected that an SR that deals with the effectiveness of a given intervention for the elderly covers only empirical studies with this age group; therefore, this is a typical case in which the inclusion criterion is justified by the context of the SR.

The inclusion and exclusion criteria should refer to the parameters established by the researchers to perform the SR. These parameters should reflect decisions to better respond to the scientific problem raised by the SR and, therefore, should not reflect limitations of the researchers. The adopted criteria should give priority to the selection of publications relevant to the objective, and not limit the selection of relevant publications (e.g. restrictions on language and year of publication, unless duly justified, are limitations to the SR rather than criteria themselves). It is suggested in this topic that the authors create a standardized form (i.e., protocol) with the detailed inclusion and exclusion criteria, trying to predict the most the variations to the rules and decisions to be made, for use in each evaluation performed by the referees in the material selection.

Extraction and data management

In this topic, the procedures and tools used to capture and manage the information from the publications found and selected should be chosen and reported. In other words, the surveyors should explain how the publications were grouped and managed. For example, researchers can rely on and report the use of software tools for the management of publications (e.g., Mendeley, Zotero, EndNote). In addition, information extracted from publications should also be informed. If the SR is followed by a meta-analysis, it is suggested that the researchers point out which information was extracted for the SR and which for the meta-analysis.

In management software tools, it is possible to group references (and abstracts, if desired) from various databases. One of the most commonly used free software tool today is Mendeley, and EndNote is reported as being the best-evaluated paid tool in the scientific context today. In both softwares, the data can be organized into folders in the most appropriate way for the evaluator, in order to make the process of assessment of references more practical and to reduce errors (e.g., omissions, deletions). In addition, these tools allow the identification and deletion of duplicates (i.e. articles found on more than one basis) for control and material analysis without redundancy. With respect to redundancy, it is worth mentioning that these tools only allow the identification of redundancy in the title and abstract, and that the evaluating referees must be attentive to check for the possibility of redundant content (e.g., selection of thesis and article extracted from the same thesis).

Evaluation and selection procedures

It is recommended that, after insertion of the publications found in the information management tool (i.e., identification) and excluded the duplicates, the collected material be screened and selected by at least two independent researchers. Each researcher must initially read the titles and abstracts of each of the publications found, preferably separating between “ineligible” (publications that do not fit the scope of the SR for sure) and “eligible” (publications that may be selected) publications. At this point it is preferable for the surveyors to give priority to the false positive, i.e., if they are unsure, they should fit the publication into the “eligible” category, so that later, with detailed reading, the reference can be properly included or excluded. The completion of this stage, called screening, occurs with the junction of the screenings carried out by the two surveyors, so that discrepancies are discussed until they reach a consensus (i.e., double consensus, or by a third referee who may occasionally be called in to settle evaluation discrepancies).

After the initial screening, publications characterized as “eligible” should be read in full and independently by the two researchers, initiating the stage called eligibility. As the final product of this stage, each of the researchers will have the set of publications selected for the SR. Again the discrepancies should be discussed among the researchers until they reach a consensus (with double or third opinion). At this point, with the final set of publications selected, the authors should be clear about the data that will be extracted from the reviews and presented as SR results and have them properly listed (e.g., year of publication, information about methodological design, main results).

At this point, the author will find the total number of publications that will actually generate data about the scope of the review. There is not a minimum of publications and data established in the literature to perform an SR, as long as it includes more than one unit, otherwise it would be a mere report of a previous publication. There are also in the literature the so-called empty reviews, which consist of reviews in which no publication met the inclusion criteria (Lang, Edwards, & Fleischer, 2007). Although more unusual, they are found in the literature – 376 empty reviews were found in the Cochrane database, according to Yaffe, Montgomery, Hopewell, and Shepard (2012) –, and allow researchers to identify gaps in certain areas of study. However, regarding the meta-analysis, at least two different data sets are required (Borenstein, Hedges, Higgins, & Rothstein, 2009). There is evidence that two publications are enough (Turner, Bird, & Higgins, 2013), but the limitations and weightings concerning the generalization need to be presented. Currently, meta-analyses display an average inclusion of three studies, but more than 30% tend to have only two studies (Davey, Turner, Clarke, & Higgins, 2011).

Assessment of quality of the report, risk of bias, and publication bias

The evaluation and verification of the quality of the report and the method used in the selected publications in an SR are fundamental parts for characterization and quality of this type of study. The assessment of the quality of the selected material may result in the weighting, or even the exclusion, of some studies at the time of description of the results and discussing the findings. This consideration or exclusion can be performed because a study with poor methodological quality and high risk of bias should be considered with great caution in the amount of pooled data.

The assessment of the quality of the report refers to the verification of the information to be presented in the scientific publication of an SR, in order to guarantee its robustness. That is, researchers should assess whether the necessary information is reported throughout the whole structure of the SR (i.e., Introduction, Method, Results, and Discussion). This assessment, therefore, is more related to the information presented and less to the quality of the procedures used. Several checklists are published for this purpose (Collins et al., 2015), so that researchers should decide which is most appropriate for the SR to be performed.

The bias risk assessment, unlike the previous one, refers specifically to the methodological procedures adopted, that is, to the quality of the procedures used in the SR. Therefore, an evaluation verifies if the information is presented and the other one verifies the quality of the methods used in the conduction of the SR (and, if appropriate, the meta-analysis). Checklists for bias risk verification (Moons et al., 2014).

In addition to the presentation of the checklists used, it is necessary that the surveyors report the scoring system that was applied, with the appropriate cut-off points, when applicable. In case the researchers made changes to the checklist, presentation of the details is required.

On the other hand, the evaluation of the publication bias is associated to the verification of trends in publications, that is, to verify if in a given area, studies are being published more according to the results (e.g., significant or insignificant) and less by the quality and impact of these studies (Rothstein, Sutton, & Borenstein, 2006). It is, then, verified whether there is evidence of favoritism in the publication for certain

types of results. One of the most commonly used procedures for this purpose is the funnel plot, along with the Egger's regression test. It should be emphasized that, in some publications, the description of the procedures for evaluation of publication bias is presented through statistical analysis.

Statistical analysis

An SR does not generally involve in-depth statistical analysis, except for the presentation of the description of the raised variables (e.g., frequency, percentage). In case of SRs followed by meta-analysis, this topic should be addressed, informing which analyses and statistical criteria were used specifically in the meta-analysis. For example, the meta-analytic model used, the types of effect sizes used, definition of the significance criteria, and the statistical programs used to conduct the analyses.

Results

Just like the different modalities of empirical studies, after the presentation of the Method, the results of the data analysis must be organized and reported. Typically, they are presented in the following sequence: flow diagram, table on the assessment of the quality of the report of the publications selected, table on the assessment of the risk of bias of the publications selected, table with general characteristics of the publications selected, and table with detailed information on the method and results of the publications selected. In addition, for the cases in which meta-analyses are reported, the graphs forest plot and funnel plot can also be shown.

However, the overall presentation of the results will also depend on how the authors worked the data and there is no predefined formula or rule. It is important that the results be explored in the most complete and clear way to achieve the study goal. Thus, sometimes it may be interesting to divide the results into subtopics, and sometimes a single topic seems to better integrate the data. Another possibility is to use the analyses of the quality of the report and risk of bias to split this section between studies with better quality and less risk of bias, and studies with more fragile data. In the following paragraphs we will briefly describe the typical information regarding the results frequently reported in SRs.

Flow chart

It is a figure in which it is graphically reported the step-by-step of the survey of publications, from the initial group of studies found to the final selection of eligible publications. It is recommended to use the PRISMA flow diagram (Moher et al., 2009), which can be downloaded from the site, in an editable and free format. As recommended in the PRISMA website, authors should cite a scientific paper (Moher et al., 2009) when presenting the PRISMA flow diagram. The flow diagram is provided as shown in Figure 1.

The diagram should be filled in properly, according to the results found in each of the steps. Two details should be clarified: first, in the upper right-hand corner ("Additional records") - authors can present, for example, publications found via Google Scholar; and second, in the last frame on the right ("Full-text"), the authors should highlight the main justifications for the exclusion of studies, using brief expressions, which can be detailed in the paragraphs.

Table: Assessment of the quality of the report and evaluation of the risk of bias of the selected publications

The presentation of results regarding the assessment of the quality of the report and risk of bias can be done in different ways, including tables, figures or the text itself. The use of figures for this purpose is

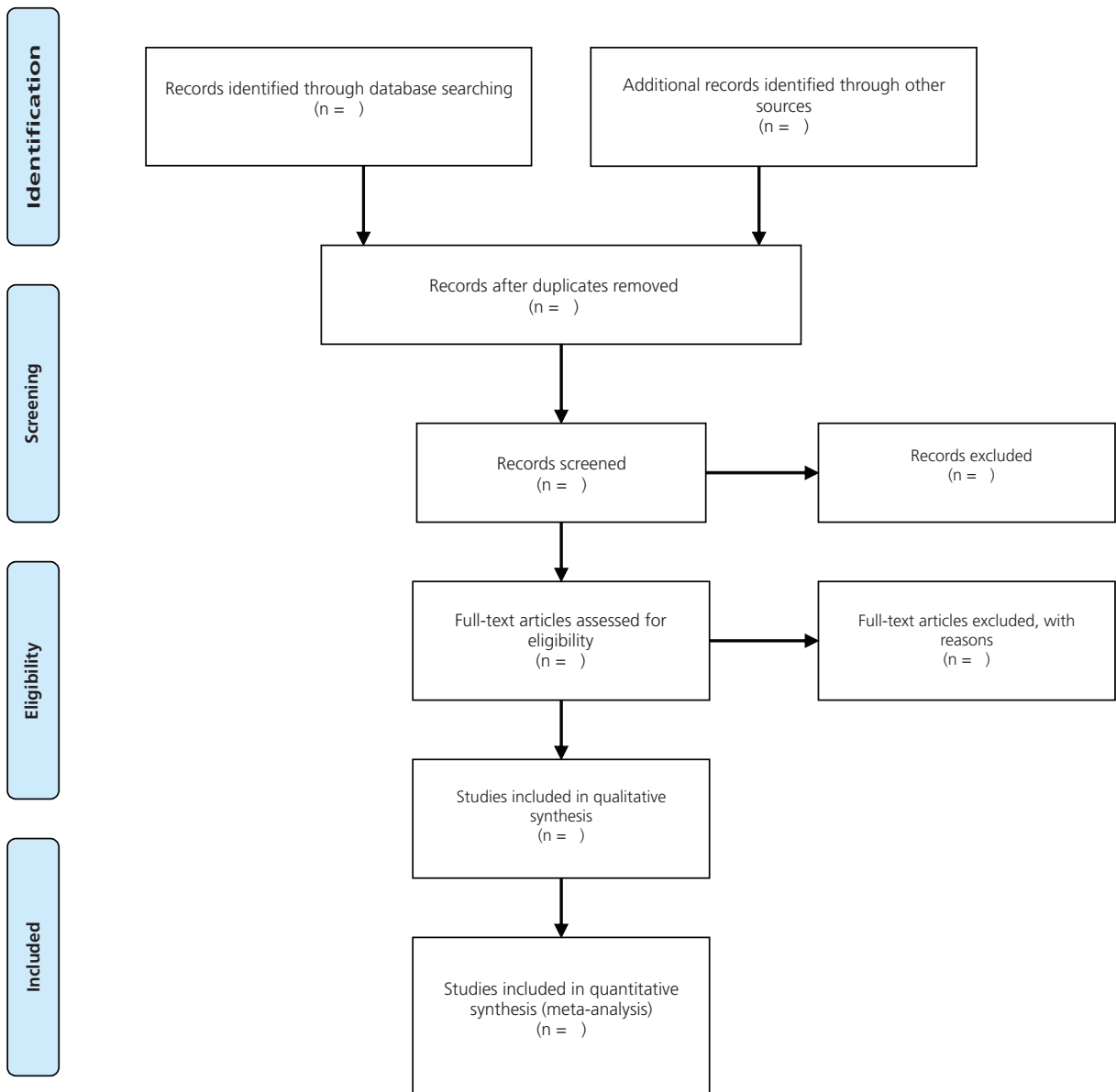


Figure 1. Model for flow diagram according to PRISMA (Moher et al., 2009).

associated with the use of programs that generate these figures automatically based on the information entered (e.g., RevMan); however, many of these programs cannot be used for any checklist because they need some information which, when not inserted, do not allow the generation of pictures. In the case of tables, it is possible to present a single table integrating the results of the quality of the report and the risk of bias, or the presentation can be performed in separate tables.

Tables: General characteristics and detailed information on method and results of selected publications

The table containing the general characteristics exhibits information which is less related to the content of the publications and more pertinent to the publication itself, such as number of authors, year of

publication, journal, impact factor, area of publication, country, among others. In general, these results are used for the understanding of the publication profile. For example, the year in which more publications were found, in which country, where more studies were conducted and in what area of science. More complex profiles can also be investigated by aggregating more than one information at the same time, for example, to note that in a given period a given country published more and, in another period, another country. Checking the profile of the publications can help to understand both the results obtained and to establish guidelines for future publications.

In terms of the table with details of the method and results, information on the method (e.g., sample, measures used, collection procedures) and information on the results obtained in the selected publications should be presented (e.g., reported effects, direct conclusions of the analyses presented). In general, the information presented in this table is, in fact, that with greater relevance for an SR in relation to the scientific problem investigated. It is in this table that the information that responds to the scientific problem posed by the review is presented.

In the paragraphs related to the two tables, summaries of the presented information should be reported, seeking to discriminate patterns between the data and allowing the reader to have access to the main results that directly respond to the objective of the SR. The search for consensus and divergences are relevant at this moment, and it will allow the discussion and conclusion of the SR to be complete and enlightening for the scientific literature in the area. Details of specific studies should be presented only when the authors consider relevant the presentation of such information.

Figures: Forest plot and funnel plot

The forest plot is the graphical representation typically exhibited in cases of meta-analysis. This is an effect estimated based on the effects of the selected studies. A forest plot is shown in Figure 2.

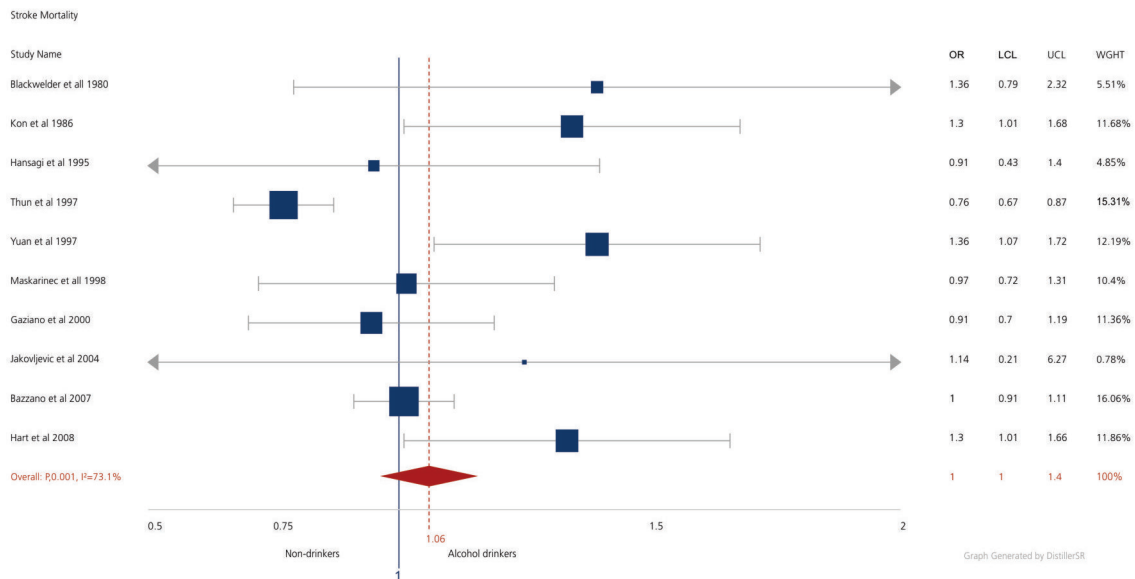


Figure 2. Hypothetical Forest plot.

Note: Distiller Systematic Review (SR) Forest Plot Generator from Evidence Partners.

The left side of the figure presents the studies, represented by the name of the authors and year of publication; in the last line, information on heterogeneity is considered. In the center, the effects of each study (i.e., squares) are presented along with their relevant confidence intervals (i.e., lines that cut the squares), as well as the diamond representing the total effect of the studies, which is the main result of the meta-analysis. On the right side there is information about sample size, effect (in this case, represented by Cohen's d) and their respective confidence intervals. Details of the information provided by forest plot are beyond the scope of this study.

The funnel plot refers to a scatter plot that, as previously explained, is used to verify the possibility of publication bias of the selected studies for a meta-analysis. Figure 3 presents a hypothetical funnel plot.

We can observe an inverted pyramid, so that publications with significant effects lie within this pyramid. Different shades ranging from white to dark gray indicate different ranges at the level of significance. Ideally, the studies are scattered to form an inverted pyramid, suggesting that there is no publication bias. However, not finding this format does not necessarily indicate the publication bias, also because with a small number of studies it is more difficult to establish this format.

Discussion

The discussion of the results in the SR report also reflects the structure typically used in publications of empirical studies in general. Sometimes the objective is resumed, and the main result obtained with the SR, considering its objective, is presented. This result is then discussed in light of the literature, outlined in the

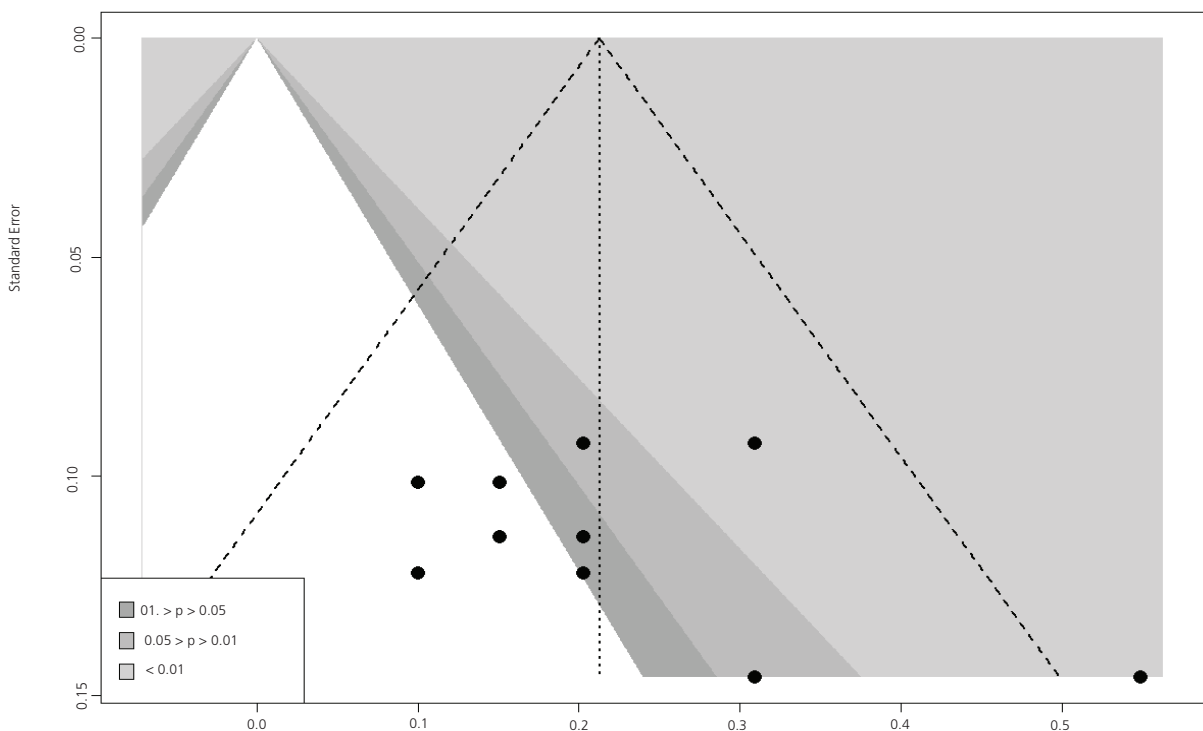


Figure 3. Hypothetical Funnel plot.

Note: Distiller SR Forest Plot Generator from Evidence Partners.

Introduction of the publication. However, in the case of the review, it is not uncommon for the Introduction to cite studies that are in the review study. It is important that the author develops the discussion with a focus on the objective, but thinking about the gaps, divergences, omissions, contradictions, answers and challenges that his review may point out. Thus, the discussion paragraphs should briefly return to the most relevant results, and this part should be used mainly to discuss these results with the support of the literature.

When structuring the discussion, the authors should note that one of the main objectives of an SR is to identify the current panorama of the survey problem. With it, the author must first respond directly to the research problem raised, making use of the information to propose directions for future investigations and to suggest new studies. The information presented in an SR will often be used to guide that area of study, to develop and improve public policies, to correct assessment and intervention measures, among other possible applications.

Final considerations

The part destined to the final considerations should allow the reader to have an information preview of the SR focus, the results achieved and their impact on the improvement of science in the area. Furthermore, transpositions between the constructed knowledge and the applied practice in the area are often exhibited. In this part, the author can express him/herself in terms of expectations that were achieved with the results and discussions already described. The finalization of the manuscript also includes the presentation and reflection about the main limitations of the SR performed, considering future possibilities to deal with these restrictions. Authors may choose to present such considerations in the topic of the discussion or in a separate topic (which will also depend on the publication standards of the journal for which they intend to submit the manuscript).

Conclusion

This article aimed to present a practical and informative guide to systematic reviews in psychology, contributing to the systematization of the procedures and methodological steps, in line with the recent inputs of the literature. Specifically, within the national framework, we intend to increase the interest and investment of researchers in SR, providing them with information to improve the quality of systematic reviews in psychology in Brazil.

Contributors

All authors participated in all stages of manuscript preparation.

References

- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York: John Wiley & Sons.
- Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMC Medicine*, *13*(1), 1. <http://dx.doi.org/10.1186/s12916-014-0241-z>

- Davey, J., Turner, R. M., Clarke, M. J., & Higgins, J. P. (2011). Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: A cross-sectional, descriptive analysis. *BMC Medical Research Methodology*, 11(1), 160. <http://dx.doi.org/10.1186/1471-2288-11-160>
- De Winter, J. C., Zadpoor, A. A., & Dodou, D. (2014). The expansion of Google Scholar versus Web of Science: A longitudinal study. *Scientometrics*, 98(2), 1547-1565. <http://dx.doi.org/10.1007/s11192-013-1089-2>
- Gehanno, J. F., Rollin, L., & Darmoni, S. (2013). Is the coverage of Google Scholar enough to be used alone for systematic reviews. *BMC Medical Informatics and Decision Making*, 13(1), 7. <http://dx.doi.org/10.1186/1472-6947-13-7>
- Grimshaw, J. (2004). So what has the Cochrane Collaboration ever done for us? A report card on the first 10 years. *Canadian Medical Association Journal*, 171(7), 747-749. <http://dx.doi.org/10.1503/cmaj.1041255>.
- Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *Plos One*, 10(9), e0138237. <http://dx.doi.org/10.1371/journal.pone.0138237>
- Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0*. Londres: The Cochrane. Retrieved February 1, 2019, from <http://handbook.cochrane.org>
- Holly, C., Salmond, S., & Saimbert, M. (2012). *Comprehensive systematic review for advanced nursing practice*. New York: Springer.
- Knipschild, P. (1994). Systematic reviews: Some examples. *British Medical Journal*, 309(6956), 719-721.
- Lang, A., Edwards, N., & Fleiszer, A. (2007). Empty systematic reviews: Hidden perils and lessons learned. *Journal of Clinical Epidemiology*, 60(6), 595-597. <http://dx.doi.org/10.1016/j.jclinepi.2007.01.005>
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. Oxônia: Oxford University Press.
- McInnes, M. D., Moher, D., Thombs, B. D., McGrath, T. A., Bossuyt, P. M., Clifford, T., ... Hunt, H. A. (2018). Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA Statement. *Jama*, 319(4), 388-396. <http://dx.doi.org/10.1001/jama.2017.19163>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Plos Medicine*, 6(7), e1000097. <http://dx.doi.org/10.1371/journal.pmed.1000097>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. <http://dx.doi.org/10.1186/2046-4053-4-1>
- Moons, K. G., Groot, J. A., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., ... Collins, G. S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *Plos Medicine*, 11(10), e1001744. <http://dx.doi.org/10.1371/journal.pmed.1001744>
- Nichols, H. (1891). The psychology of time. *The American Journal of Psychology*, 3(4), 453-529. <http://dx.doi.org/10.2307/1412061>
- Pauling, L. (1987). *How to live longer and feel better*. New York: Avon Books.
- Petticrew, M. (2001). Systematic reviews from astronomy to zoology: Myths and misconceptions. *British Medical Journal*, 322(7278), 98-101. <http://dx.doi.org/10.1136/bmj.322.7278.98>
- Petticrew, M., & Roberts, H. (2006). How to appraise the studies: An introduction to assessing study quality. In M. Petticrew & H. Roberts (Eds.), *Systematic reviews in the social sciences: A practical guide* (pp.125-163). <http://dx.doi.org/10.1002/9780470754887.ch5>
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. New Jersey: John Wiley & Sons.
- Treasure, E. T., Chestnutt, I. G., Whiting, P., McDonagh, M., Wilson, P., & Kleijnen, J. (2002). The York review, a systematic review of public water fluoridation: A commentary. *British Dental Journal*, 192(9), 495-497. <http://dx.doi.org/10.1038/sj.bdj.4801410>
- Turner, R. M., Bird, S. M., & Higgins, J. P. (2013). The impact of study size on meta-analyses: Examination of underpowered studies in Cochrane reviews. *Plos One*, 8(3), e59202. <http://dx.doi.org/10.1371/journal.pone.0059202>
- Yaffe, J., Montgomery, P., Hopewell, S., & Shepard, L. D. (2012). Empty reviews: A description and consideration of Cochrane systematic reviews with no included studies. *Plos One*, 7(5), e36626. <http://dx.doi.org/10.1371/journal.pone.0036626>
- Zoltowski, A. P. C., Costa, A. B., Teixeira, M. A. P., & Koller, S. H. (2014). Qualidade metodológica das revisões sistemáticas em periódicos de psicologia brasileiros. *Psicologia: Teoria e Pesquisa*, 30(1), 97-104.