

**The research context**

A huge amount of research is produced each year, often with conflicting findings. These between-study differences may be due to study differences, flaws or chance (sampling variation). In such situations, it is not always clear what the overall picture is, or which results are most reliable and should be used as the basis for practice and policy decisions.

**SYSTEMATIC REVIEW**

Systematic reviews aim to address these problems by *identifying, critically evaluating* and *integrating* the findings of all relevant, high-quality individual studies addressing one or more research questions. A good systematic review might achieve most or all of the following (Baumeister & Leary, 1997; Bem, 1995; Cooper, 2003):

- Establish to what extent existing research has progressed towards clarifying a particular problem;
- Identify relations, contradictions, gaps, and inconsistencies in the literature, and explore reasons for these (e.g. by proposing a new conceptualisation or theory which accounts for the inconsistency);
- Formulate general statements or an overarching conceptualization (make a point, rather than summarizing all the points everyone else has made; Sternberg, 1991);
- Comment on, evaluate, extend, or develop theory;
- In doing these things, provide implications for practice and policy;
- Describe directions for future research.

A systematic review is therefore a piece of research in its own right and, by its nature, is able to address much broader questions than single empirical studies ever can (e.g. uncovering connections among many empirical findings; Baumeister & Leary, 1997). Indeed, systematic reviews sit above all other research designs at the top of the 'hierarchy of evidence' because they have the potential to provide the most important practical implications.

**Systematic reviews are characterised by being *objective, systematic, transparent and replicable*.** They involve a systematic search process to locate studies which address a particular research question, as well as a systematic presentation and synthesis of the characteristics and findings of the results of this search. The criteria for inclusion and exclusion in the review are objective, explicitly stated and consistently implemented such that the decision to include or exclude particular studies is clear to readers and another researcher using the same criteria would likely make the same decision. This explicit approach aims to minimise bias and allows readers of the review to assess the author's assumptions, procedures, evidence and conclusions, rather than taking the author's conclusions on faith. This methodology also allows other researchers to update the review at a later time in order to integrate new findings. **To best-achieve the purposes of a systematic review, Baumeister (2013) recommends adopting the mindset of a judge and jury, rather than a lawyer.** A judge and jury sceptically evaluate the evidence in order to render the fairest judgment possible. In contrast, a lawyer's approach to the evidence involves trying to make the best case for one side of the argument.

**Systematic review definition:** 'A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies' (Cochrane Collaboration, 2014).

**THIS GUIDE DESCRIBES THE KEY STAGES IN CONDUCTING A SYSTEMATIC REVIEW****(I) SCOPING****(IV) SCREENING****(II) PLANNING****(V) ELIGIBILITY****(III) IDENTIFICATION (SEARCHING)****IT THEN COVERS:****(VI) DECIDING WHEN TO DO A QUANTITATIVE OR A QUALITATIVE RESEARCH SYNTHESIS****(VII) HOW TO PRESENT (WRITE UP) A SYSTEMATIC REVIEW**

**(I) SCOPING****(1). Formulate one or more research questions**

What do you want to know and about what topics? Who will your audience be? Do you have a clear idea of the type of research finding that will be relevant to addressing your research question(s)? Clear, specific and answerable research question(s) are essential to a successful review. For instance 'Is family therapy an effective treatment for anorexia nervosa?' will be a more effective question and produce a more focused set of search results than 'How do I help people with eating disorders?' You will probably need to do some 'scoping' of the literature to find out what has been done before and what might make a novel, important and interesting scientific contribution to the literature. Examining a narrow research question or research area will make your task much simpler, faster and easier!

**(2). Thoroughly clarify whether the planned systematic review has already been done**

To avoid wasting time and energy, establish whether a systematic review of your research question(s) has already been done, or is registered as an ongoing review. Search thoroughly! This search will begin to familiarise you with the literature, save you time if a review already exists and does not need updating, or help provide a rationale for why you are conducting an updated review. Some example rationales for conducting an updated rather than an original systematic review might be: (i) it has been 10 years since the last systematic review on this topic and the literature has rapidly expanded since the last review, meaning new studies need to be accounted for; (ii) the last review was methodologically flawed in various ways (e.g. it was not systematic), which you intend to address with your review; (iii) the last review focused on X but you think it's worth focusing on Y for particular theoretical or empirical reasons. If a review has already been done and does not need repeating (i.e. the existing review addresses the same question(s), is of adequate quality, and is relatively recent), that review will form useful reading for your project.

**(II) PLANNING****(3). Break your research question(s) down into individual concepts to create search terms**

Search terms operationalize your research question(s) and help you find as many potentially relevant articles as possible to include. You are aiming to conduct a search that is exhaustive and therefore representative of all studies that have been conducted on the topic of interest. Reading the existing literature and talking to your supervisor will give you a good idea of how to translate your research question(s) into search terms.

It will be worth thinking of alternative terms and concepts which address the same question as it is common for a range of terms to be used to describe the same phenomenon or research area (this will come from reading and knowing the literature). Consider:

- Synonyms (e.g. 'recycle', 'refuse', 'salvage', 'recover')
- Singular/plural forms, verbal forms, adjectives (e.g. 'recycling' 'recycled')
- Different spellings (e.g. 'colour', 'color')
- Broader/narrower terms (e.g. 'Britain', 'England', 'Scotland', 'Wales')
- Classification terms used by databases to sort their contents into categories listed by headings and sub-headings, if relevant to your search

There is a balance between sensitivity (finding as many articles as possible that may be relevant) and specificity (making sure those articles are indeed relevant); at this stage, you want your search terms to err on the side of sensitivity so that you do not miss anything.

**(4). Formulate preliminary inclusion and exclusion criteria - and then review these in the initial stages of the literature searching and sifting process**

Based on your knowledge (and your supervisor's knowledge) of the literature, formulate a list of objective (unbiased) inclusion and exclusion criteria which will allow you to address the research question(s) you are posing, ensure the quality and similarity of included studies, and clearly define the boundaries of the review. Apply these inclusion and exclusion criteria consistently throughout the review

process. Studies that are eligible for inclusion will meet the inclusion criteria and not meet the exclusion criteria. The particular inclusion and exclusion criteria used will depend on the topic of the systematic review, as well as theoretical, empirical and methodological issues specific to that literature. The broader the research domain, the more detailed inclusion and exclusion criteria tend to become (Lipsey & Wilson, 2001). You will need to justify your decision to adopt particular inclusion and exclusion criteria based on theory and evidence. It goes without saying that your reasons for including particular studies need to be objective, rather than, for instance, disagreeing with a particular author's conclusions or not liking a particular author. Readers of the review will interpret the results and conclusions of the review within the context of these inclusion and exclusion criteria (e.g., generalizability, practical implications).

Common inclusion and exclusion criteria concern:

- Research question (topic, scope);
- Definition or conceptualisation (terms and concepts are often defined differently, depending on theoretical or empirical considerations. For example, interventions such as 'stress management' and 'relaxation training' may be defined differently by different authors);
- Measures/key variables (what was measured; how e.g., self-report; did measures need to meet particular psychometric criteria to be included);
- Research design (e.g., observational studies, experimental studies, quantitative studies, qualitative studies);
- Participants (e.g., adult, child, business leaders);
- Time frame (e.g. since the start of the literature, since the last review);
- Data (e.g. if doing a meta-analysis, studies need to report an effect size on the relationship of interest or else provide sufficient information that could be used to compute this statistic).

For each particular research area, researchers have to consider whether it is better to include a larger number of studies (a fuller representation of the available research on a particular topic; potentially broader and more generalizable conclusions), or a more focused group of studies (which may improve quality and the confidence that can be placed in conclusions).

Best practice is of course to formulate inclusion and exclusion criteria *before* you begin the review. However, as someone who is not already a complete expert on the subject area, this will be difficult or impossible, so personally I think this criteria can be flexed slightly, but only in terms of coming up with *which* criteria to run with. However, this slight flexibility is inappropriate in relation to how the inclusion and exclusion criteria are *applied* during literature searching and sifting. (When you write up the review, you would simply say that you used particular inclusion and exclusion criteria). So the most likely process for undertaking a systematic review initially (until you are clear on the focus) will probably be a cycling between coming up with appropriate inclusion and exclusion criteria, and objectively and systematically applying these during literature searching.

When you begin sifting the literature, you will encounter particular studies that are 'near misses' or 'borderline cases' – studies which could be almost equally argued for inclusion or exclusion, and which require more of a subjective judgment call. These studies require careful consideration, recourse to existing theory and evidence, and probably discussion and shared decision-making between you and your supervisor. Your inclusion and exclusion criteria may need to be revised based on the conceptual issues these studies raise. If you revise your inclusion and exclusion criteria after having started searching and sifting the literature, you may well need to repeat the literature searching and sifting process from the start again to ensure that excluded studies which meet this new criterion are now included.

#### **(5). Create clear record keeping systems and keep meticulous records by working systematically**

Before actually doing any searching and sifting of the literature, it is worth creating one or more record keeping systems to record what you do and your decision-making. This may seem an unnecessary effort but if the literature is large (or massive), it is almost impossible to remember exactly what you did and

why in relation to hundreds of different decisions over months or years. If you need to repeat or check anything, this record will save you a lot of time.

I suggest:

- Make a record of the details of the searches you do and the results;
- Make a list of the number of studies excluded at the screening stage (i.e. based on their Title and/or Abstract)
- Make a table to record the list of studies excluded at the potential eligibility stage (i.e. based on reading the full-text), along with reasons for excluding each study based on your inclusion and exclusion criteria. Common reasons for exclusion are: publication type (e.g., non-research article); study design (e.g., unsuitable data); measure (e.g., unvalidated measure(s); participants (e.g., youth sample); other aspects of the inclusion and exclusion criteria). This step is particularly important because it justifies why studies that people might have expected to have been included, were excluded;
- Make a table describing the efforts made to find and retrieve unpublished work.

There are a number of guidelines which outline how to report systematic reviews and meta-analyses, including many which are methodology-specific (CONSORT guidelines for reporting randomised controlled trials). The **AMSTAR** (Assessment of Multiple Systematic Reviews) and **PRISMA** (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklists are very good and applicable across different research areas. **Best practice for systematic review is to present a flow diagram of the literature searching and sifting process (e.g., The PRISMA flow diagram is very good; see References at end). As a minimum, you will need to describe everything you did and why in detail.**

### **(III) IDENTIFICATION (SEARCHING)**

#### **(6). Use your search terms to search at least two different (relevant) electronic databases**

Having reached this stage, the aim is now to find all available published and unpublished work which addresses your research question(s), operationalized through your search terms. The best way to find the vast majority of published work which addresses your research question(s), is to comprehensively search at least two different electronic databases:

- Select databases that are relevant to your topic area (e.g. Medline, EMBASE, ISI Web of Knowledge);
- Consider which parts of articles you want to search (e.g. abstract, full text, title);
- Consider using limits and filters within particular databases to search by article type (e.g. review articles and research syntheses, empirical articles), subject categories, sub-headings, etc;
- Consider using Boolean search operators to broaden or narrow your search: 'AND' (a search for all of your search terms; e.g., 'heart AND lung'), 'OR' (a search for at least one of your search terms; e.g., 'heart OR lung'), 'NOT' (a search to exclude certain search terms; e.g., 'heart not lung');
- Consider using a truncation symbol to look for all words starting with a particular combination of letters (use '\$' or '\*' depending on the database) (e.g. 'dep\$' or 'dep\*' will retrieve 'depression', 'depressive', 'depressed', etc);
- Consider using a wildcard symbol to stand-in for one character, if that character exists ('#' or '?' or '\$' depending on the database) (e.g. 'wom#n' will find 'women' or 'woman');
- Consider using the truncation symbol (\*) between words to match words (e.g. 'midsummer \* dream' will return results that contain the exact phrase, 'a midsummer night's dream');
- Consider using parentheses because commands within these run first (e.g. '(smok\* OR tobacco)')
- Consider searching by proximity, to search for one word within *n* number of words of another word ('adjn' or 'NEAR/n' depending on the database) (e.g. 'patient adj3 anxiety' will retrieve records where 'patient' and 'anxiety' appear within three words of each other);
- Consider excluding unwanted meanings (e.g. 'jaguar NOT(car OR cars)', if you want the animal);

- Consider searching by publication year, if you have a good theoretical or empirical reason for doing so (e.g. 2001-current if you are conducting a review since a previous review which was published in 2001);

Best practice guidelines for conducting systematic reviews suggest that the literature search and sifting process is ideally done by two separate reviewers who must both agree on work to be included. The process for resolving disagreements between assessors is specified explicitly in the review (e.g., disagreements over inclusion are discussed and, where possible, resolved by consensus after referring to the inclusion and exclusion criteria and relevant theoretical and empirical issues). However, as this is part of a PhD, you will probably be conducting the majority or all parts of the systematic review yourself.

### **(7). Carefully inspect the search results**

Examine the search results and read a few of the better quality and more recent, relevant articles.

- Do the search results suggest that your inclusion and exclusion criteria are reliable and effective in identifying potentially relevant articles and balancing specificity and sensitivity?
- If not, do you need to revise your inclusion and exclusion criteria, or search terms?
- Do the search results reveal new search terms that would make a useful addition to your existing search terms?

**If your search results suggest that you need to modify your search terms and/or inclusion and exclusion criteria, return to the planning stage (II), and re-run the search.** For example, a search for a study on dietary interventions during pregnancy failed to find a particular trial because a particular team of authors called a dietary intervention a 'lifestyle intervention' (and there was no subject-index heading in the database for 'diet').

### **(8). Conduct additional searches to ensure you have located all potentially relevant published and unpublished work**

Database searching may miss some potentially relevant work, so additional searches are required for published and unpublished work.

#### *Published work*

There are a number of methods which may uncover potentially eligible published (and perhaps unpublished) work that may have been missed at the database searching stage:

- Read the Reference section of work that is suitable for inclusion, which was located through database searching; this will provide you with: (i) a list of potentially relevant work, and (ii) a list of journals which contain a number of relevant studies, which you can then hand-search if needed;
- If needed, hand-search (manually search) journals. This can sometimes identify articles (and other work, e.g. letters) which have not yet been included in electronic databases, and those which are not indexed or which have been indexed incorrectly;
- If your previous database search (e.g. Medline and Web of Science) did not cover non-English-language articles and conference proceedings, re-search for these;
- Locate relevant book chapters, perhaps for data, but more likely for references of relevant work that you want to track down;
- It may be the case that the information reported in a published study is insufficient to make a decision about inclusion. In such a circumstance, you would try to contact the author to enquire about additional details or data.

#### *Unpublished work*

Systematic reviews aim to be as comprehensive and representative of the literature they describe as possible. As such, one key aspect of the methodology of systematic reviews is a concerted effort to search for and include relevant unpublished work (that meets the inclusion and exclusion criteria) in order to reduce the effects of **publication bias**. Publication bias (also called the 'file drawer problem;' Rosenthal, 1979) and 'bias against the null hypothesis;' Cooper, 2003) describes the tendency for the

availability of published research to depend on the results (Begg, 1994; Vevea & Woods, 2005). Studies have historically been selectively published dependent on the statistical significance of the results because statistical significance has been considered good/important, whilst statistical nonsignificance has been considered bad/trivial. However,  $p$ -values are limited because they incorporate information about the size of an effect and the size of the sample. Therefore:

- A statistically significant  $p$ -value may reflect a large effect size, or a small effect size that has been measured in a large study;
- A statistically nonsignificant  $p$ -value may reflect a small effect size, or a large effect size measured in a small study;
- Two studies with exactly the same effect sizes could vary greatly in their significance level depending upon the number of participants (see Cumming, 2014).

Publication bias therefore occurs when studies that found statistically nonsignificant findings do not submit their results for publication, or if they do, their manuscript is rejected for publication by reviewers and/or journal editors (Borenstein et al., 2009; Lipsey & Wilson, 2001). Publication bias poses a potentially serious threat to the validity of the conclusions of a systematic review: if research is published depending on the statistical-significance of the results, it is likely that published studies will have more 'positive' results (larger effect sizes) than unpublished studies. Therefore, if systematic reviews include only published studies, this will result in an inflated impression of the literature (and in meta-analysis, an over-estimation of population effect sizes) (Lipsey & Wilson, 2001). A biased impression of the literature could potentially lead to inappropriate conclusions being drawn, with potentially serious practical implications (Lipsey & Wilson, 2001). Although awareness of the flaws of null hypothesis significance testing (NHST) is becoming more and more prominent (which may reduce publication bias), the use of NHST will endure for the foreseeable future because it is the dominant zeitgeist and because of the allure of dichotomous statistically significant/nonsignificant decision-making (see Cumming, 2014).

"A reason frequently given for excluding unpublished research (from systematic reviews) is that it is often of lesser quality than published research...this is too simple a dichotomy. For example, researchers often do not publish their results because publication is not their objective. Some research associated with degree requirements is conducted by individuals who do not seek academic careers. Other research is conducted as evaluations for agencies making internal decisions about program effectiveness. The decision to publish (or let work perish) is not isomorphic with judgments about quality... Conversely, most researchers would agree that some low-quality research does get published. Moreover, research is often turned down for publication for reasons other than quality. In particular, research that fails to achieve standard levels of statistical significance." (Cooper, 2003, p.6). For these reasons, it is now accepted practice that rigorous research syntheses include both published and unpublished research (Cooper, 2003; Lipsey & Wilson, 2001).

Unpublished, potentially relevant work can be located in a number of ways. The most important of these involves contacting researchers with one or more publication on the topic for forthcoming data, or further details of existing data. Decide in advance how much time will be given to allow authors to reply, balancing consideration of their busy workload with your own needs. For example, you might adopt a one-month cutoff through: initially contacting authors for information (explaining who you are and why you are contacting them); waiting two weeks; sending a nice follow-up email if they do not reply; waiting another two weeks; then stopping.

Depending on the research question and topic, you may need to search for 'grey literature,' which is any literature produced in electronic or print format that has not been controlled by commercial publishers. Examples include technical or research reports from government agencies, reports and working papers from scientific research groups and committees, doctoral dissertations, conference proceedings, official publications, etc.

Different databases specialise in technical reports, unpublished research, the work of government departments, conference proceedings, and theses. Some examples:

- opengrey.eu is a European database compiled by different national libraries in various European countries who submit any grey literature they receive;
- opendoar.org/Directory of Open-Access Repositories website searches the open-access repositories of thousands of universities all over the world;
- WorldCat database for dissertations and theses;
- Google and Google Scholar are reasonably effective in locating dissertations and work by societies and charities.

#### **(IV) SCREENING**

##### **(9). Export references to a citation manager to collate the search results**

Your search probably revealed a large number of results. Exporting search results to a citation manager (e.g., EndNote, RefWorks) has a number of advantages:

- It will save a massive amount of time as it is an electronic rather than a manual process;
- Your search results are saved and backed up, meaning that you cannot lose this valuable information;
- The citation manager can identify and get rid of duplicate versions of the same work;
- You can obtain, and share with others, full-text versions of many of the articles identified;
- The citation manager will compile your reference list and format it in an array of referencing styles (this can be very useful if you submit to a particular journal which uses a particular referencing style, the manuscript is rejected, and then you decide to submit to another journal which uses a different referencing style).

##### **(10). Read the Title and/or Abstract of identified work**

Read the Title and/or Abstract of all work identified by your searches. Most work will not meet your inclusion and exclusion criteria. If Title and/or Abstract suggest that the work is potentially eligible for inclusion (i.e. may pass your inclusion and exclusion criteria), you will need to obtain the full-text version and read that carefully. At this stage, continue to err on the side of sensitivity (locating and sifting as many articles as possible) so that you do not miss anything. For record keeping purposes during the screening stage, it is sufficient to make a list of the number of rejected articles (rather than noting the reasons for excluding each study).

#### **(V) ELIGIBILITY**

##### **(11). Sift the full-text version of potentially eligible articles and extract relevant information to be included**

Your focus now shifts from sensitivity to specificity (carefully making sure potentially eligible studies are indeed relevant and appropriate for inclusion); you now need to sift the full-text version of potentially eligible articles to see if each is indeed suitable for inclusion. Even during this stage, you can rapidly reduce the pool of potential studies by focusing your reading on whether or not each published or unpublished work meets your inclusion and exclusion criteria (i.e. focus on the Method and Results sections, rather than the Introduction and Discussion). Keep a record of why each piece of published or unpublished work was rejected during this stage (i.e. how the work failed to meet particular inclusion criteria), as this increases the transparency of the selection process.

When you identify work for inclusion, you need to carefully and thoroughly extract all relevant information for potential inclusion (what is considered relevant information will depend on your research questions(s) and topic). The information you extract will predominantly relate to your inclusion and exclusion criteria and therefore will likely cover: definition or conceptualisation; measures/key variables; research design; participants; year of publication; data/results; study design, setting, etc.

Your inclusion and exclusion criteria are designed to ensure that high-quality relevant work is included. However, you may additionally think that it would be constructive to make notes on the quality of each

included study, as it will be helpful to summarise methodological limitations which bias the literature (if you are doing a qualitative synthesis), and/or to examine whether study quality systematically biases effect sizes (if you are doing a quantitative synthesis). There is no consensus on the best way to assess study quality, but most methods encompass issues such as: appropriateness of study design for addressing the research objectives; participant or condition selection methods; measurement of study variables; control of confounding; appropriate use of statistics; quality of reporting; quality of intervention/condition; generalizability; author conflict of interest (see the EQUATOR Network: [www.equatornetwork.org/](http://www.equatornetwork.org/); but also see Greenland & O'Rourke, 2001, for a discussion of study quality issues in meta-analysis).

### **(VI) DECIDING WHEN TO DO A QUANTITATIVE OR A QUALITATIVE RESEARCH SYNTHESIS**

Whether a qualitative or a quantitative approach to research synthesis is most appropriate will depend on the nature of the data you have, your research question(s), the aims of your review, and theoretical and empirical issues.

#### **Meta-analysis**

A quantitative systematic review is conducted by performing a meta-analysis. Meta-analysis is concerned with estimation; questions such as "How large is the effect?" or "To what extent . . . ?" For example, "Does playing violent computer games influence aggressive behaviour?" or "Is a treatment, intervention or innovation effective, and is it more effective than a comparable approach?" Broad categories of research questions where meta-analysis may be applicable are:

- Central tendency research (e.g., prevalence rates)
- Pre-post contrasts (e.g., growth rates)
- Group contrasts
  - Experimentally created groups (e.g., comparison of outcomes between treatment and comparison groups)
  - Naturally occurring groups (e.g., comparison of spatial abilities between boys and girls)
- Association between variables
  - Measurement research (e.g., validity generalization)
  - Individual differences research (e.g., correlation between personality constructs) (Lipsey & Wilson, 2001)

Meta-analysis numerically synthesises the available evidence for a given question. Meta-analysis would be appropriate when a collection of studies:

- Report quantitative results, rather than qualitative findings, or theory;
- Examine the same or similar constructs/relationships and are therefore meaningfully comparable;
- Derived from similar research designs (e.g., it would not be appropriate to combine experimental trials of treatments for depression, with observational studies in which the level of depression was correlated with the level of service received – although, in this case, separate meta-analyses might be appropriate);
- Results are bivariate/zero order relationships/single-degree of freedom contrasts (i.e. the relationship between two variables), rather than relationships that have been adjusted for the effect of additional variables (e.g., partial or multivariate effects);
- Have results that can be configured as standardised effect sizes (Borenstein et al., 2009; Lipsey & Wilson, 2001).

A qualitative research synthesis would be appropriate when:

- A group of studies are so methodologically diverse as to make meta-analytic aggregation impractical;
  - Conceptual and methodological approaches to research on a topic have changed over time and you argue that there is benefit in reviewing all of this research;
  - When developing a new theory or critiquing one or more existing theories;
  - When reviewing measurement approaches in a particular literature
- (see Baumeister, 2013; Baumeister & Leary, 1997 for further, detailed discussion of narrative reviews).

**(VII) HOW TO PRESENT (WRITE UP) A SYSTEMATIC REVIEW**

As a general rule, if you (i) adhere to respected guidelines which outline how to report systematic reviews and meta-analyses (e.g., AMSTAR, PRISMA), (ii) present a flow diagram of the literature searching and sifting process (e.g., PRISMA flow diagram), and (iii) use as a template for formatting and content, at least two different systematic reviews which have been published in top journals in your field, and which address similar goals to the purpose of your review, you can't go far wrong!

There is no one right way to do a review; what you cover and how you organize the review should be dictated by the goals you have for the review. For example, reviews could be arranged historically, so that topics are introduced in the chronological order in which they appear in the literature; conceptually, so that works relating to the same ideas appear together; or methodologically, so that works that use similar methods are grouped together (see Cooper, 2003, p. 5, for further details).

More specifically, a systematic review should probably include most or all of the following, depending on whether you are doing a qualitative or a quantitative review:

**Introduction**

- Provide a theoretical and empirical background to the literature, explaining key terms, definitions and concepts;
- Provide a theoretical and/or empirical rationale for the systematic review;
- Explicitly state what the “foci, goals, perspective, coverage, organization, and audiences are...This statement should appear early in the review so that the reader can construct an appropriate frame of reference for evaluating the effort” (Cooper, 2003, p. 5). “Few readers can manage to wade through 50 pages of text...and findings before learning what the point is” (Baumeister & Leary, 1997, p. 316). This can be achieved either by (i) presenting an existing or new theoretical conceptualization at the start of the review and then using the remainder of the manuscript to review the literature relevant to the theory, or by (ii) presenting a quick summary of an existing or new theoretical conceptualization at the start of the review, postponing its full elaboration until after the literature has been reviewed (Baumeister & Leary, 1997).
- Describe aims and objectives;

**Method**

- Describe a comprehensive, objective, systematic literature search in detail, including how and when particular databases were searched; years searched; search terms; inclusion and exclusion criteria with a theoretical and/or empirical rationale; what concerted efforts you made to locate and include all published and unpublished work on the topic (i.e. what comprehensive and systematic preventative steps were taken to minimize bias and errors in the study selection process);
- Discuss ‘borderline’ or ‘near-miss’ cases which were excluded and which readers might have expected to see included, or which were included despite partially breaching the inclusion and exclusion criteria (see Trickey et al., 2012);
- Provide a flow diagram (e.g. PRISMA) describing identification, screening, eligibility and inclusion stages, with number of studies included and excluded at each stage, along with reasons for exclusion during the eligibility (full-text versions of articles) stage;

**Results**

- Describe the characteristics of included studies in detail in a Table;
- Describe assessment of the scientific quality of included studies;
- Critically evaluate and integrate the results in an unbiased and systematic way, including stating whether the results are heterogeneous and discussing possible reasons for this (if studies were somehow pooled, explain why it was appropriate and meaningful to do so);
- Balance conflict resolution (by identifying inconsistencies in study results) against bridge building (by identifying points of contention in theories, conceptualizations, and methods in the literature) (Cooper, 2003).

Discussion

- Summarise and discuss the findings and conclusions of the review in a balanced and impartial way, in the context of previous theory, evidence and practice;
- Explicitly and intuitively link your conclusions to the evidence reviewed;
- Discuss the strengths and limitations of the literature and, by implication, the review, including considering the scientific quality of included studies and methodological problems in the literature (e.g., methodological rigor or lack thereof, the amount of evidence, its consistency, and its methodological diversity). Conclusions should be tempered by the flaws and weaknesses in the evidence. Perhaps propose a new conceptualisation or theory which accounts for inconsistencies. (Baumeister & Leary, 1997);
- Establish to what extent existing research has progressed towards clarifying a particular problem/formulate general statements or an overarching conceptualization. Quantitative or qualitative reviews may conclude that the available evidence suggests one of four possibilities (see Baumeister & Leary, 1997, for a detailed discussion of these):
  - 1) A hypothesis is correct, at least based on the present evidence
  - 2) A hypothesis, although not proven, is currently the best guess and should be assumed to be true until contrary evidence emerges
  - 3) It is not clear whether a hypothesis is true or false
  - 4) A hypothesis is false
- Comment on, evaluate, extend, or develop theory;
- Draw conclusions and make recommendations for practice;
- Describe directions for future theory, evidence and practice by pointing out remaining unresolved issues (Baumeister & Leary, 1997).

Appendix

- Consider including a detailed list of studies excluded at the potentially eligible (full-text versions of articles) stage;
- For the purposes of a thesis, you will probably also want to include sample record keeping forms and your completed records.

Some general points about how to present a narrative review

“The literature reviewer should first ensure that he or she has covered the research accurately and thoroughly. But literature reviewers should also ask themselves whether they have presented each study in a way that makes its relation to the integrative themes clear and explicit” (Baumeister & Leary, 1997, p. 317). Therefore, **describe existing findings in the context of providing something new to the literature** (e.g. a novel theoretical contribution, providing a summary and critique of the current literature, providing directions for future research) (Baumeister & Leary, 1997). This is best achieved by structuring the review into sub-sections, with a coherent “story” and flow throughout. **“The reader needs to know how strong or how weak the overall evidence for each main point is. Group or section critiques accomplish this better than criticizing each individual study.** The most useful form of critique is normally the following: after describing the methods and results of a group of studies relevant to some point, the author should indicate briefly the major flaws in the methods and what alternative explanations they raise. Next, the consistency of the findings should be considered. Then the author should assess the quantity and especially the methodological diversity of the evidence, keeping in mind that consistency across large quantities of methodologically diverse evidence is the best available substitute for having the proverbial direct pipeline to the truth. Finally, the author should provide a summary as to how strong the evidence is” (Baumeister & Leary, 1997, p. 318).

**Provide a balanced coverage and critique of all aspects of the literature, ideally spending some time searching for evidence that would seem to contradict the main conclusions and patterns.** These can be included in the article as separate sections and given the same critical appraisal as the supporting evidence. If there are important exceptions to the general patterns and conclusions, the literature review is strengthened by acknowledging them, and theory can be built further by recognizing moderators and

boundary conditions. If the exceptions are merely apparent and do not on close inspection contradict the main pattern, the manuscript is strengthened by pointing them out" (Baumeister & Leary, 1997, p. 319).

**Cite a study's conclusion whilst also describing the method and specific results to provide a context** (Baumeister & Leary, 1997). For example "in a sample of A, method B produced result C (Reference), thereby supporting the view that X causes Y" (Baumeister & Leary, 1997, p. 317).

References:

- Baumeister, R. F. (2013). Writing a literature review. In M. J. Prinstein & M. D. Patterson (Eds.), *The portable mentor: Expert guide to a successful career in psychology* (pp. 119-132; 2<sup>nd</sup> ed.). New York: Springer Science+ Business Media.
- Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology, 3*, 311-320.
- Bem, D. J. (1995). Writing a review article for Psychological Bulletin. *Psychological Bulletin, 118*, 172-177.
- Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399-409). New York: Russell Sage Foundation.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.
- Cooper, H. M. (2003). Editorial. *Psychological Bulletin, 129*, 3-9.
- Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2008). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29.
- Field, A. P. & Gillett, R. (2010). How to do meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*, 665-694.
- Greenland, S., & O'Rourke, K. (2001). On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics, 2*, 463-471.
- Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. London: Sage Publications.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. *British Medical Journal, 339*, 332-336.
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology, 7*, 10-17.
- Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., Henry, D. A., & Boers, M. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology, 62*, 1013-1020.
- Sternberg, R. J. (1991). Editorial. *Psychological Bulletin, 109*, 3-4
- Trickey, D., Siddaway, A.P. Meiser-Stedman, R., Serpell, L., & Field, A.P. (2012). A meta-analysis of risk factors for post-traumatic stress disorder in children and adolescents. *Clinical Psychology Review, 32*, 122-138.
- Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*, 428-443.

Recommended reading for literature reviews, systematic reviews and meta-analysis:

**Basic texts**

- Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine, 5*, 101-117.

A very basic overview of writing literature reviews which is focused towards biomedicine, but is very accessible and covers lots of things that are so basic they are not covered in the other recommended reading texts.

- Randolph, J. J. (2009). A guide to writing the dissertation literature review. *Practical Assessment, Research & Evaluation, 14*, 1-13.

Again, covers the very basics. Repeats some of Cooper (2003).

- Sternberg, R. J. (2000). Article writing 101: A crib sheet for 50 tips for the final exam. In R. J. Sternberg (Ed.), *Guide to publishing in psychology journals* (pp. 199-205). Cambridge, UK: Cambridge University Press.

Broad range of very brief tips, covering different aspects of article writing and the publication process.

- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review, 4*, 1-12.

Covers basic concepts for different aspects of writing a literature review.

### **More sophisticated texts**

#### **Narrative reviews:**

- Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology, 3*, 311-320.

Comprehensive and erudite overview of problems, advantages, opportunities and pitfalls of different types of narrative literature review.

- Bem, D. J. (1995). Writing a review article for Psychological Bulletin. *Psychological Bulletin, 118*, 172-177.

Detailed guidance on how to write a comprehensive narrative review article; also briefly covering the journal revision process.

- Cooper, H. M. (2003). Editorial. *Psychological Bulletin, 129*, 3-9.

Detailed overview of different types of quantitative and qualitative literature reviews, including presenting a taxonomy to help you get your head around how and when to do different types of reviews.

#### **Meta-analyses:**

- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.

A totally comprehensive, very accessible "how to guide" to all aspects of meta-analysis.

- Field, A. P. & Gillett, R. (2010). How to do meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*, 665-694.

As the name suggests, another (very good) "how to guide."

- Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. London: Sage Publications.

Very clearly written and accessible; if it's not in Borenstein et al (2009), try here.

#### **Additional reading:**

- Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2008). *The handbook of research synthesis and meta-analysis* (2<sup>nd</sup> ed.). New York: Russell Sage Foundation.

Detailed, scholarly, highly specialist chapters on most aspects of research synthesis and meta-analysis.

- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29.

Comprehensive, scholarly and very clear, detailed discussion of problems with null-hypothesis statistics testing ( $p$ -values), and arguments for using estimation instead (i.e. effect sizes, confidence intervals, meta-analysis).

- Kline, R. B. (chapter 4, 2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: APA Books.

Summary of deep flaws in null-hypothesis statistics testing, including mistaken beliefs, damaging practises and ways in which null-hypothesis statistics testing hinders research progress.

- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, 59-82.

Detailed discussion of the advantages and disadvantages of meta-analysis, effect size metrics and very brief coverage of how to do a meta-analysis.

- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118, 183–92.

Discusses how to present meta-analyses. A little dated now but still very useful and clear; compare Rosenthal's suggestions with example published meta-analyses in your target journal.