3

Searching the Literature

After articulating one or more research questions for your meta-analysis (Chapter 2), the next step is to locate the studies that will provide information to answer these questions (as described in subsequent chapters on coding and analysis). Unlike narrative reviews that are typically unsystematic in their searching of the literature (or at least typically do not articulate this process), the field of meta-analysis has devoted considerable attention to practices of searching and retrieving relevant literature.

In this chapter, I describe how it is useful to conceptualize the studies in your meta-analysis as a sample of a larger population (Section 3.1) and how this conceptualization leads to explicit criteria of which type of studies should be included versus excluded from your meta-analysis (Section 3.2). I will then describe various methods of searching for relevant literature, considering the advantages and disadvantages of each (Section 3.3). I conclude the chapter by describing the importance of "reality checking" your search (Section 3.4) and the practical matter of creating a meta-analytic database (Section 3.5). The steps involved in a literature search as described in this chapter are summarized in Figure 3.1.

3.1 DEVELOPING AND ARTICULATING A SAMPLING FRAME

Given that meta-analysis uses the individual study as its unit of analysis, it is useful to think of your meta-analysis as consisting of a sample of studies, just as primary analyses sample people or other units (e.g., families, businesses)

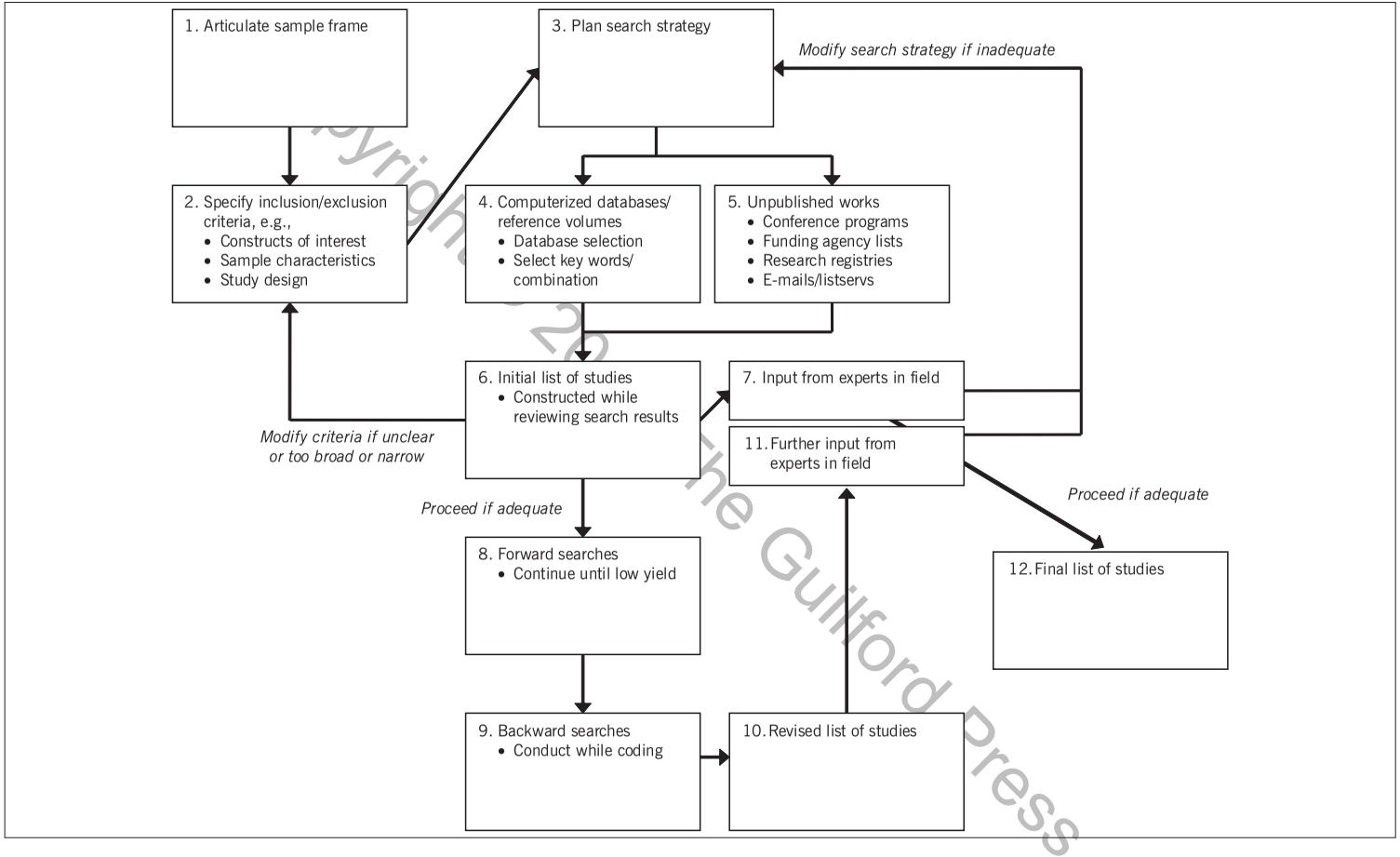

FIGURE 3.1. Basic steps of searching the literature.

comprising its sample. In primary analyses, we typically wish to make inferences to a larger population that is represented by the sampled individuals; in meta-analysis, we typically wish to make inferences to a larger population of possible studies from the sample of studies included in our review. In both cases, we want our sample to be representative of this larger population, as opposed to a biased (nonrepresentative) set.

To illustrate the importance of obtaining an unbiased sample of studies, we can consider the threat of publication bias (discussed in further detail in Chapter 11). The top of Figure 3.2 displays a hypothetical population of effect sizes, with the horizontal (*x*) axis representing the effect sizes obtained in studies of this population and the vertical (*y*) axis representing the frequency that studies yield this effect size.[1] We see that the mean effect size in this population is somewhere around 0.20 and that there is a certain amount of deviation around this mean due to either sampling fluctuation or unspecified (random) differences. The bottom part of this figure shows the distribution of a biased sample of studies drawn from this population. I have used arrows of different width to represent the likelihood of studies from the population being included in this sample. The arrows to the right are thick to represent studies with large effect sizes being very likely to be included in the sample (i.e., very likely to be found in a search), whereas the arrows to the left are thin to represent studies with small effect sizes being very unlikely to be included in the sample (i.e., likely not found in a search). We can see that this differential likelihood of inclusion by effect sizes results in a biased sample. If you were to meta-analyze studies from this sample, you would find a mean effect size somewhere around 0.30 rather than the 0.20 found in the population. Thus, analysis of this biased sample of studies leads to biased results in a meta-analysis.

The goal of searching and retrieving the literature for a meta-analytic review is to obtain a representative, unbiased collection of studies from which inferences can be made about a larger population of studies. Meta-analyses differ from primary analyses in that your goal is typically to obtain *all* of the studies comprising this population as it currently exists.[2] Whether or not you are successful in obtaining all available studies (and it is not possible to know with certainty that you have), it is still appropriate to consider this set of studies as a sample, from which you might draw inferences about a larger population including studies you did not locate or studies performed in the future (assuming that these studies are part of the same population as those included in your meta-analysis).

This approach, in which you think of the studies included in your meta-analysis as a sample from a population to which you wish to make inferences,

has two important implications. First, this conceptualization properly frames the conclusions you draw from results after completing your meta-analysis; this is important in allowing you to avoid either understating or overstating the generalizability of your findings. Second, and more relevant during the planning stages of your review, this conceptualization should guide your criteria for which type of studies should or should not be included in your meta-analysis, as described next.

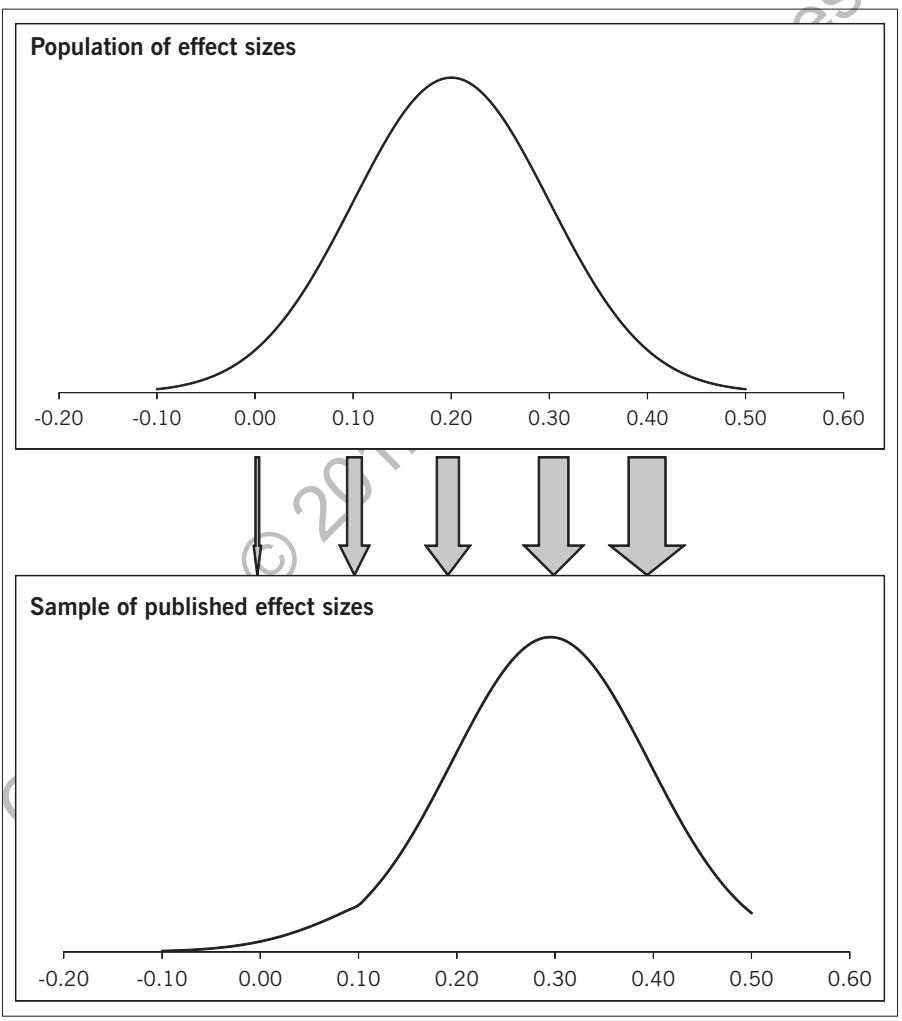

FIGURE 3.2. Hypothetical illustration of biased sample due to differential likelihood of including studies in a meta-analysis.

3.2 INCLUSION AND EXCLUSION CRITERIA

The inclusion criteria, and conversely the exclusion criteria, are a set of explicit statements about the features of studies that will or will not (respectively) be included in your meta-analysis. Ideally, you should specify these criteria before searching the literature so that you can then determine whether each study identified in your search should be included in your meta-analysis. Practically speaking, however, you are likely to find studies that are ambiguous given your initial criteria, so you will need to modify these criteria as these unanticipated types of studies arise.

3.2.1 The Importance of Clear Criteria

Developing an explicit set of inclusion and exclusion criteria is important for three reasons. First, as I noted earlier, these criteria should reliably guide which studies you will (or will not) include in your meta-analysis. This guidance is especially important if others are assisting in your search. Even if you are conducting the search alone, however, these criteria can reduce subjectivity that might be introduced if the criteria are ambiguous.

The second reason that explicit criteria are important is that these criteria define the population to which you can make conclusions. A statement of exclusion (i.e., an exclusion criterion) limits your conclusions *not* to involve this characteristic. For example, in the example meta-analysis I will present throughout this book (considering various effects involving relational aggression), my colleagues and I excluded samples with an average age of 18 years or older. It would therefore be inappropriate to attempt to draw any conclusions regarding adults from this meta-analysis. A statement of inclusion (i.e., an inclusion criterion) implies that the population is defined—at least in part—by this criterion. For example, a criterion specifying that included studies must use experimental manipulation with double-blind procedures would mean that the population is of studies with this design (and any other inclusion criteria stated).

The third reason that explicit criteria are important relates to the goal of transparency, which is an important general characteristic to consider when reporting your meta-analysis (see Chapter 13). Here, I mean that your inclusion/exclusion criteria should be so explicit that a reader could, after performing the same searches as you perform, come to the same conclusions regarding which studies should be included in your meta-analysis. To illustrate, imagine that you perform a series of searches that identify 100 studies, and based on your inclusion/exclusion criteria you decide that 60

should be included in your meta-analysis. If another person were to evaluate those same 100 studies using your inclusion/exclusion criteria, he or she should—if your criteria are explicit enough—identify the same 60 studies as appropriate for the review. To achieve this level of transparency in your meta-analysis, it is important to record and report the full set of inclusion/exclusion criteria you used.

3.2.2 Potential Inclusion/Exclusion Criteria

The exact inclusion/exclusion criteria you choose for your meta-analysis should be based on the goals of your review (i.e., What type of studies do you want to make conclusions about?) and your knowledge of the field. Nevertheless, there are several common elements that you should consider when developing your inclusion/exclusion criteria (from Lipsey & Wilson, 2001, pp. 18–23):

3.2.2.a Definitions of Constructs of Interest

The most important data in meta-analyses are effect sizes, which typically are some index of an association between *X* and *Y*.[3] In any meta-analysis of these effect sizes, it is important to specify criteria involving operational definitions of both constructs *X* and *Y*. Although it is tempting for those with expertise in the area to take an "I know it when I see it" approach, this approach is inadequate for the reader and for deciding which studies should be included. One challenge is that the literature often refers to the same (or similar enough) construct by different names (e.g., in the example meta-analysis, the construct I refer to as "relational aggression" is also called "social aggression," "indirect aggression," and "covert aggression"). A second challenge is that the literature sometimes refers to different constructs with the same name (e.g., in the example meta-analysis, several studies used a scale of "indirect aggression" that included such aspects as diffuse anger and resentment that were inconsistent with the more behavioral definition of interest). By providing a clear operational definition of the constructs of interest, you can avoid ambiguities due to these challenges.

3.2.2.b Sample Characteristics

It is also important to consider the samples used in the primary studies that you will want to include or exclude. Here, numerous possibilities may or may not be relevant to your review, and may or may not appear in the literature you

consider. Some basic demographic variables to consider include gender (e.g., Will you include studies sampling only males or only females?), ethnicity (e.g., Will you include only representative samples, or those that sample one ethnic group exclusively?), and age (e.g., Will you include studies sampling infants, toddlers, children, adolescents, young adults, and/or older adults?). It is also worth considering what cultures or nationalities will be included. Even if you place no restrictions on nationality, you will need to exclude reports written in languages you do not know,[4] which likely precludes many studies of samples from many areas of the world. Beyond these examples, you might encounter countless others—for example, samples drawn from unique settings (e.g., detention facilities, psychiatric hospitals, bars), selected using atypical screening procedures (e.g., certain personality types), or based on atypical recruitment strategies (e.g., participants navigating to a website). Although it is useful if you can anticipate some of these irregular sample characteristics in advance, many will invariably arise unexpectedly and you will have to deal with these on a case-by-case basis.

3.2.2.c Study Design

A third consideration for inclusion/exclusion criteria for almost every meta-analysis is the type of research design that included studies should have. Some obvious possibilities are to include only experimental, quasi-experimental, longitudinal naturalistic, or concurrent naturalistic designs. Even within these categories, however, there are innumerable possibilities. For example, if you are considering only experimental treatment studies, should you include only those with a certain type of control group, only those using blinded procedures, and so on? Among quasi-experimental studies, are you interested only in between-group comparisons or pre–post designs? Answers to these sorts of questions must come from your knowledge of the field in which you are performing the review, as well as your own goals for the meta-analysis.

3.2.2.d Time Frame

The period of time from which you will draw studies is a consideration that may or may not be relevant to your meta-analysis. By "period of time," I mean the year in which the primary study was conducted, for which you might use the proxy variable year of publication (or completion, presentation, etc., for unpublished works). For many phenomena, it might be of more interest to include studies from a broad range of time and evaluate historic effects through moderator analyses (i.e., testing whether effect sizes vary regularly

across time; see Chapter 9) rather than a priori excluding studies. However, in some situations it may make sense to include only those studies performed within a certain time period. These situations might include when you are only interested in a phenomenon after some historic changes (e.g., correlates of unprotected sex after the AIDS crisis) or when the phenomenon has only existed during a certain period of time (e.g., studies of cyberbullying have only been performed since the popularity of the Internet has increased).

3.2.2.e Publication Type

The reporting format of the studies is another consideration for potential inclusion/exclusion criteria. Although including only published studies is generally considered problematic (due to the high possibility of publication bias; see Chapter 11), it is important to consider what types of reports will be included. Possibilities include dissertations, other unpublished written reports (e.g., reports to funding agencies), conference presentations, or information that the researcher provides you upon request.

3.2.2.f Effect Size Information

Finally, a necessary inclusion criterion is that the studies provide sufficient information to compute an effect size.[5] In most situations, this will be information provided in the written report that allows you directly to compute an effect size (see Chapter 5). However, you should also consider whether you would include studies that provide only enough information to compute a lower-bound estimate (e.g., probability ranges such as $p < .05$, statements that results were nonsignificant; see Chapter 5). When studies do not report sufficient information to compute effect sizes, you should contact the study authors to request more information; here, a necessary inclusion criterion is that the authors supply this information.

3.2.3 Relative Advantages of Broad versus Narrow Inclusion Criteria

In developing inclusion/exclusion criteria, specifying both broad and narrow sets of criteria has notable advantages. By broad criteria, I refer to a set of criteria that include most possible studies and exclude few, whereas narrow criteria will exclude many studies and include few. Of course, these two choices represent end points along a continuum. Selecting a set of criteria that falls along this continuum has several implications for your meta-analysis.

Perhaps the most important consideration in weighing a broad versus narrow set of criteria is that of the population of studies about which you want to draw conclusions. Put simply: Would you prefer to make conclusions about a very specific, well-defined population, or would you rather make more generalizable conclusions about a potentially messy population (i.e., one with likely fuzzy boundaries, likely inconsistent representation in your sample of studies, and possibly undistinguished subpopulations)? Specific to the issue of study quality (see Chapter 4) is the question of whether you want to include only the most methodologically rigorous studies or are willing to include methodologically flawed studies (risking the "garbage in, garbage out" criticism described in Chapter 2). There is not a universal "right answer" to these questions, just as there is not a right answer to the issue of level of generalization to the "apples and oranges" problem described in Chapter 2. If you choose a narrow set of criteria, you should be cautious to draw conclusions only about this narrowly defined population. In contrast, if you choose a broad set of criteria, it is probably advisable to code for study characteristics that contribute to this breadth and to evaluate these as potential moderators of effect sizes (see Chapter 9).

A second consideration is the number of studies that will ultimately be included in your meta-analysis by specifying a broad versus narrow set of criteria. Broad criteria will result in a meta-analysis of more studies that are more diverse in their features, whereas narrow criteria will result in fewer studies that are more similar in their features. Having fewer studies will sometimes result in inadequate power to evaluate the average effect size (see Chapters 8 and 10), will usually preclude thorough consideration of study characteristics that account for differences in effect sizes (i.e., moderator analyses; see Chapter 9), and might even lead your audience to view your review as too small to be important to the field. In contrast, having more studies increases the amount of work involved in the meta-analysis (especially the coding of studies), perhaps to the point where a meta-analysis of the full collection of studies is impossible.[6] Therefore, one consideration is to specify inclusion/exclusion criteria that yield a reasonable number of studies given your research questions, your available time and resources, and typical practices in your field. This is not the only, or even primary, consideration, but it is a realistic factor to consider.

3.3 FINDING RELEVANT LITERATURE

After specifying inclusion/exclusion criteria, the next step is to begin searching for empirical studies that fit within this sampling frame. In searching for this relevant literature, you have many options, each with advantages and

limitations over the others. Although it is not always necessary to use all of the options I list next, it is useful to consider at least most of them and how reliance on some but not others might bias the sample of studies you obtain for your meta-analysis.

Before describing these search options, it is useful to consider the concepts of recall and precision (see White, 2009). Recall is the percentage of studies retrieved from those that should be retrieved (i.e., the number of studies meeting your inclusion criteria that actually exist); it is a theoretical value that can never be known because you never know how many studies actually exist. Precision is the percentage of retrieved studies that are relevant (i.e., actually meet your inclusion criteria). Ideally, we would like both to be 100%, such that our search strategies yield every available study that meets our criteria and none that do not. In reality, we can never meet this goal, so you must balance the relative costs of one or the other being less than 100%. The cost of imperfect recall is that you will miss studies that should have been included, resulting in reduced statistical power and potentially biased results if the missed studies differ from those you included. The cost of imperfect precision is that we will waste our resources retrieving and reading studies that will not be included in our meta-analysis. Although this might not seem like a tremendous cost, it is if it means that you cannot complete your meta-analysis.[7] The goal of your search strategy should be to achieve high recall without diminishing precision beyond an unacceptable level, where "unacceptable" depends on your available resources and the expected benefits of increasing recall in terms of statistical power and reducing bias.

3.3.1 Electronic Databases

Modern electronic databases, available via the Internet through most university libraries (or available for subscription for others), have made the task of searching for relevant studies much easier than in the early days of meta-analysis. Electronic databases exist in many fields, such as economics (EconLit), education (ERIC), medicine (Medline), psychology (PsycINFO), and sociology (Sociological Abstracts), to name just a few. These databases often have wide coverage (though see cautions below) and therefore serve as one of the primary search tools in modern meta-analysis. In fact, these databases are typically the first searches performed by meta-analysts, and I would consider them necessary (though not sufficient) for your meta-analysis.

Despite their power and apparent simplicity, using electronic databases is a more complex process than might be initially apparent (see Reed & Baxter, 2009). I next describe three considerations in using these databases, attempting to consider these generically rather than focusing on any one database.

3.3.1.a What Is Included and What Is Excluded?

The first question you should ask before using any electronic database is "What is included (and what is excluded) from this database?." Answering this question requires you to read the documentation of the databases you are considering; consulting with librarians in your topical area is invaluable, as they have considerable expertise on this question.

Some databases include dissertations and other unpublished works, whereas others do not. If the database you plan to use does not include dissertations, you should certainly supplement your search of this database with one that includes dissertations (such as Proquest dissertation and thesis database). If the database does not include other unpublished work, and your inclusion criteria allow for this work, then you will need to ensure that other search strategies will find these works. If the database does include unpublished works, you should investigate how these works are selected for inclusion; databases that include works unsystematically (e.g., primary study authors being willing to submit works to the database) should be treated cautiously as the sample of unpublished work may be biased.[8]

Another consideration is the breadth of published work included in the database. Prominent journals are more likely to be included than peripheral journals, and books by larger publishers are more likely to be included than those by lesser-known publishers. If it is plausible that the results (effect sizes) could differ in studies published in outlets included (e.g., prominent journals) versus excluded (e.g., periphery journals) in the database(s) you are using, then reliance on this database may yield a biased sample of studies.

3.3.1.b Key Words

After researching the databases you will use to understand their coverage, you then search the databases for relevant studies. To perform this search, you generally enter key words, for which the search engine will return records containing these key words. Selection of appropriate key words goes far in increasing recall and precision, so you should consider these key words carefully and report them in your meta-analytic review.

A first consideration is the key words you select. You can select key words based on your knowledge of the literature in your area, by examining the key words specified in studies that you know contain data about the phenomenon of interest, and through thesauri available in some electronic databases. Your goal is to create a list of words or phrases that (1) are as specific to the phenomenon you are investigating as possible and (2) cover the range

of terms used to describe the phenomenon. Considering the example meta-analysis involving associations of relational aggression with various other constructs (e.g., gender, peer rejection), our goal was to search for all studies of relational aggression. Terms such as "aggression" were too broad, as these would identify studies investigating constructs aside from that in which we were interested. Using the term "relational aggression" was more specific, but by itself would have been inadequate because different researchers use different terms for this construct. We ultimately developed a list of four terms to use in our search ("relational aggression," "social aggression," "indirect aggression," and "covert aggression") that represent the terms typically used by primary study authors investigating this construct.

Wildcard marks (e.g., "*" in PsycINFO) are useful in combination with key words. Wildcard marks are used in conjunction with a stem, specifying that the search engine returns all studies containing the specified stem followed by any characters where the wildcard mark is typed. For example, submitting the phrase "relational agg*" would return studies containing the phrases "relational aggression," "relational aggressor," and so on. Using wildcard marks can also return unexpected and unwanted findings, however, (e.g., the example stem and wildcard would also return any studies that used the phrase "relational aggravation"). These can generally be recognized quickly and skipped, or you can modify the wildcard search term or use the Boolean statement "not" as described next.

Boolean statements are a tremendous asset when you are searching electronic databases. These statements include "or," "and," and "not" in most databases. The use of "or" is especially valuable in combining alternative key words for the same construct; for example, we connected the four terms for the construct of interest using "or" in our example meta-analysis (i.e., the search phrase was: "relational aggression" or "social aggression" or "indirect aggression" or "covert aggression"). The logical statement "and" is useful for either limiting the studies returned or specifying two construct associations that are of interest in many meta-analyses. For example, in the example meta-analysis, we could have combined the above search (various key words for relational aggression combined using "or") with a phrase limiting the samples to childhood or adolescence ("child* or adolesc*") using the "and" statement.[9] Similarly, if we were only interested in studies reporting associations between relational aggression and peer rejection (one of the examples I use commonly throughout the book), we could have used "and" to link the phrases for relational aggression with a set of phrases for peer rejection. Finally, you can use the key word "not" either to exclude unwanted wildcard phrases (e.g., in the example above, I could specify "not 'relational aggrava-

tion'" to remove the unwanted studies using this term), or to specify exclusion criteria (e.g., specifying "not 'adult'").

3.3.1.c Cautions

Electronic databases are incredibly powerful and time-efficient tools for searching for relevant studies, and I believe that every modern meta-analysis should use these databases. However, at least three cautions merit consideration.

First, as I described earlier, you should carefully consider what is *not* included in the electronic databases you use. If a database does not include (or if it has poor rates of inclusion) unpublished works or studies published in peripheral outlets, then reliance on this database alone would result in diminished recall. This diminished recall can threaten your meta-analysis by decreasing statistical power and, if the studies not included in the database systematically differ from those included (e.g., publication bias, Chapter 11), by producing biased results. To avoid these problems, you should identify alternative electronic databases and other search strategies that are likely to identify relevant studies not included in the electronic database you are using.

A related caution comes from the fact that most electronic databases are discipline specific. Although the databases vary in the extent to which they include works in related disciplines, this disciplinary specificity suggests that you should not rely on only a single database within your discipline. Many, if not most, phenomena that social scientists study are considered within multiple disciplines. For example, research on relational aggression might appear not only in psychology (e.g., in the PsycINFO database), but also in criminal justice, education, gender studies, medicine, public health, and sociology (to name just a few possibilities). I recommend that you consider searching at least one or two databases outside of your primary discipline to explore how much literature might be obtained from other disciplines.

A third caution in using electronic databases relates to their very nature: You perform a search and a list of studies is provided, but you have no idea how many potentially relevant studies were *not* provided. In other words, relying only on electronic databases provides no information about what studies were not identified in your search, so the possibility remains that some studies—and possibly even some very well-known studies—did not match your specified search criteria. You can address this problem in several ways. One possibility is to perform some additional searches within your selected database(s) that use broader terms (e.g., "aggression" rather than more spe-

cific terms such as "relational aggression") and visually scan results to see if any additional relevant studies could be identified with broader search criteria. Second, you can rely on additional search procedures besides the electronic database. I return to this topic of assessing the adequacy of your search (including the adequacy of electronic database searches) in Section 3.4.

3.3.1.d Conclusions about the Use of Electronic Databases

Electronic databases of journal articles, books and chapters, and often some unpublished works exist in most social science disciplines. These searchable databases can provide an efficient method of searching for studies to include in your meta-analysis if you carefully consider the coverage of the databases you use and select appropriate key words along with wildcard marks and Boolean statements. These electronic databases should not be your only method of searching the literature, however, as several cautions need to be considered when using them. Nevertheless, the electronic databases are likely to be one of the primary ways you will search for studies, and every modern meta-analysis should use these tools.

3.3.2 Bibliographical Reference Volumes

Bibliographical reference volumes are printed works that provide information similar to electronic databases (e.g., titles, authors, abstracts), often listing studies by broad topics and/or including an index of key words. These volumes were frequently published by large research societies and were intended to aid literature searches in specific fields in much the same way that electronic databases now do in most fields. For example, the American Psychological Association published *Psychological Abstracts* from 1927 to 2006. In many fields, publication of these printed reference volumes has been discontinued in favor of online electronic versions (though exceptions may exist).

Searching these reference volumes is not nearly as convenient as searching electronic databases, and few meta-analysts currently rely on these volumes as their primary search instrument (though you are likely to see them used when you read older meta-analytic reviews). Nevertheless, there still may be instances when you would consider using these printed volumes. Specifically, if studies potentially relevant for your meta-analysis include older studies, and the electronic databases that you use have not yet incorporated all of these older studies, then it may be useful to consult these reference volumes to ensure that you do not systematically exclude these older studies.

3.3.3 Listings of Unpublished Works

As I mentioned briefly in Chapter 2, and describe in detail in Chapter 11, one of the most challenging threats to many meta-analyses is that of publication bias (a.k.a. the "file drawer problem"). The extent to which you can avoid and evaluate this threat depends on your searching for and including unpublished studies in your meta-analysis. I have already mentioned the value of searching electronic databases that include dissertations as one method of obtaining unpublished studies. Next I list three additional listings that might allow you to find more unpublished studies. For each, I suggest searching with the same careful rigor I suggested for searching electronic databases.

3.3.3.a Conference Programs

A potentially valuable way to find unpublished studies is to search the programs of academic conferences in which relevant work is likely to be presented. Dedicated meta-analysts often have shelves of these programs, though even this idea is becoming antiquated as more conference programs are archived and searchable online. In this approach, you search the titles of presentations listed in conference books (larger conferences typically have at least crude indices) and request copies of these works from authors (whose contact information is usually listed in these books).

From my experience, it is usually possible to identify a large number of unpublished works by searching conference programs; however, retrieving copies of these presentations for coding can be more difficult. Typically, you are better able to contact authors and more likely to receive requested presentations if you make your request shortly after the conference rather than several years later. Therefore, studies obtained through conference programs probably underrepresent older studies. Some other tips I have learned through experience include: (1) whenever you request a conference presentation, provide exact details such as the title of the presentation and the year and conference where it was presented; (2) contact coauthors if you do not receive a response from the first author, as some authors of the presentation may have graduated or left academia; (3) tell the author why you are requesting this information (I will elaborate on this piece of general advice below).

Although I think conference presentations are a valuable source of unpublished studies, there are some limitations and cautions to consider. First, your search for conference presentations should be systematic. If you decide to search the programs of a particular conference, you should make reasonable efforts to search the programs' books across a reasonable number

of years (vs. the years you attended but not the years in between when you did not attend), and you should certainly search for works within the entire conference book (vs. just the presentations you happened to attend). Second, you should recognize that the response rate to your requests might be low (you should track this response rate as it might be useful to report), and you should consider the possibility that responses might be systematically related to effect sizes.[10] Finally, you should anticipate that conference presentations will often present information needed for study coding (Chapter 4) and effect size calculation (Chapters 5–7) in less detail than other formats (e.g., journal articles). It is still better to code what you can from these studies than not to consider them at all, and it is possible to request further information from study authors.

3.3.3.b Funding Agency Lists

Another valuable way to obtain unpublished studies is to search funding award listings from relevant funding agencies (e.g., National Institutes of Health, National Science Foundation, private foundations). Because funding decisions are made before results are known, studies obtained through this approach will not likely be subject to biases in findings of significance/nonsignificance. Furthermore, searching these listings is likely to yield studies that have been started but have not yet gone through the publication process (i.e., more recent studies).

3.3.3.c Research Registries

Some fields of clinical science have established listings in which researchers are expected to register a study before conducting it. To encourage registration, some journals will only publish results from studies registered prior to conducting the study. Such registries, by creating a listing of studies in advance of knowing the results, should yield a collection of results unbiased by the findings (e.g., nonsignificant or counterintuitive findings). If the field in which you are performing your meta-analysis has such registries, these will be a very valuable search avenue for obtaining an unbiased set of studies.

3.3.4 Backward Searches

After accumulating a set of studies for potential inclusion in your meta-analysis, you will begin the process of coding these studies (see Chapters

4–8). You should read these articles completely (vs. going straight for the method and results sections where most information you will code appears), searching for cited studies that might be relevant for your review that you did not identify through your other strategies. Similarly, you should carefully read prior reviews (narrative or meta-analytic) searching for potentially relevant studies.

This process of searching for relevant studies cited in the works you have found is referred to as "backward searching" (sometimes also called "footnote chasing"); that is, you are working from the studies you have "backward" in time to identify previously conducted studies cited in these works. This approach is especially useful in identifying older studies, whereas it is unlikely to identify newer studies that have not yet been cited. An important potential bias of this approach comes from the possibility that studies yielding certain "favorable" results (e.g., significant findings, effects favoring expectations) are probably more likely to be cited than studies with "unfavorable" results (e.g., null findings, counterintuitive findings).

Despite the potential biases of backward searches, I believe that they represent a valuable method of searching. My own experience is that many studies come from this approach even with what I consider quite thorough initial searches using other means. This approach is also valuable in identifying literature that might have been missed in other search approaches due to failures to use appropriate key words or to search literatures in other disciplines.

3.3.5 Forward Searches

Whereas backward searches attempt to find studies cited in the studies you have, forward searches attempt to find studies that cite the studies you have. Forward searches are often performed using special databases for this purpose (e.g., Social Science Citation Index), though some field-specific databases are incorporating this approach (e.g., the psychology database PsycINFO now has this capacity). To perform a forward search, you enter information for a study you know is relevant to the topic of your meta-analysis, and the search engine finds works that cite this study. Because these citing studies necessarily occur after the cited study, the search is moving "forward" in time and is more likely to find newer articles than a backward search.

There are various degrees of intensity in engaging in forward searches. A less intense approach is to identify several of the earliest and most seminal works on the topic, then perform forward searches to identify studies citing these seminal papers. At the other end of the spectrum, you could perform

forward searches of all works that you have determined meet the inclusion criteria for your review.

Forward searches are likely to yield high recall, as it is unlikely that many relevant studies would fail to cite at least some of the seminal works in the area. However, my experience[11] is that forward searches are often quite low in precision. This is because many papers will cite a seminal work in an area when this area is of tangential interest to the paper.

3.3.6 Communication with Researchers in the Field

The final search approach that I will describe is to consult experts/researchers in the field in which you are performing your meta-analysis. This approach actually consists of several possibilities.

At a minimum, you should ask some experts to examine your inclusion/exclusion criteria and the list of studies you have identified, requesting that they note additional studies that should have been included. If you examine these suggested studies and some do meet your inclusion criteria, then you should not only include these studies, but also consider why your search strategy failed to identify these studies (and revise your search strategy accordingly). I recommend that you consult colleagues who have a somewhat different perspective in the field than your own (i.e., different "camps") to provide a unique perspective.

Another valuable approach to communicating with researchers is simply to e-mail those individuals who conduct research in the area of your meta-analysis, asking them if they have any additional studies on the topic. This effort can also vary in intensity, ranging from e-mailing just the most active researchers in the field to e-mailing every author of studies you have identified through other means. Although you will have to identify an approach that works best for you given your field and relationships with other researchers, some practices that I have found valuable are: (1) to clearly state why I am requesting studies (e.g., "I am conducting a meta-analytic review of the associations between *X* and *Y*"); (2) to provide a small number of the most critical inclusion criteria (e.g., "I am interested in obtaining studies involving children or adolescents"); and (3) to state the various ways that they could provide the requested information to me (e.g., "I would like the correlation between *X* and *Y*, but can compute this from most other statistics you might have available, such as *t*-tests, ANOVA results, or raw means and standard deviations. I am also happy to compute this correlation myself, if you are willing to share the raw data with the agreement that I will delete this data file after computing this correlation.").

A related but less targeted approach is to post requests on listservs, webpages, or similar forums. Many of the same practices that are valuable when e-mailing are useful in such postings, though the standards of particular forums might necessitate briefer requests.

These communications with researchers are extremely valuable, though several considerations are important. First, my impression is that the response rates vary widely for different meta-analysts, with some receiving almost no responses but others receiving tremendous responses. I suspect that the factors that improve response rates include your ability to convince others that your request is important and worth their time, your ability to minimize the burden on the researchers, and the quality of relationships you have with these colleagues. A second consideration is the obvious fact that the more widespread your requests (i.e., numerous e-mails or public postings), the more people know that you are conducting this particular meta-analysis, which is a consideration in terms of the review process. Perhaps the most important consideration, however, is one that I believe means that you *absolutely must*, to at least some degree, involve colleagues in the area of your meta-analysis: Meta-analytic reviews synthesize the body of knowledge in an area of study and typically provide the foundation for the next wave of empirical study in this area. Thus, the research community has a vested interest in this process, and the meta-analyst has an obligation to consider their input. This statement does not mean that you need to send the initial draft of your meta-analysis to everyone in your field (you should not), nor that your review needs to support the conclusions of everyone in your field (your conclusions are hopefully empirically driven). Instead, by soliciting input from others in your field, whether by simply including the full body of their empirical results in your review or obtaining input from a smaller number of colleagues, your meta-analysis will benefit from this collective knowledge.

3.4 REALITY CHECKING: IS MY SEARCH ADEQUATE?

Regardless of what methods of searching the literature you rely upon, the most important question is whether your search is adequate. You can think of the adequacy of your search in three ways. First, is the sample of studies you have obtained representative of the population of studies, or is it instead biased (as illustrated in Figure 3.2)? Second, does the sample of studies you have obtained provide sufficient statistical power to evaluate the hypotheses you are interested in (or, similarly, does it provide sufficiently narrow confidence intervals of effect size estimates to be useful)? Third, would the typi-

cal scholar in my field find the sample of studies complete, or have I missed studies that obviously should be included? The first two questions directly affect the quality of the empirical conclusions of your meta-analysis and so are obviously important. The third question is less important to the conclusions drawn, but is pragmatically relevant to others' viewing of your review as adequate. This is a worthy consideration affecting both the likelihood of publication of your review and the impact it will have on your field.

The question of whether the sample you have obtained is an unbiased representation of the population is impossible to answer with certainty. However, there do exist methods of evaluating for the most likely bias—that of publication bias—which I describe in Chapter 11.

Probably the best way to answer all of these questions satisfactorily is to make every reasonable effort to ensure that your search is exhaustive—that is, to ensure that the sample of studies for your meta-analysis contains as close to all the studies that exist in the current population as possible. This goal is probably never entirely attainable, yet if you have made every effort to obtain all available studies, it is reasonable to conclude that you have come "close enough."[12] No one knows when "close enough" is adequate, and there is less empirical evidence to inform this decision than is desired, but I offer the following suggestions for your own consideration of this topic.

First, you should conduct an initial search using some combination of the methods described above that you expect will provide a reasonably thorough sample of studies. For example, you might decide to consult prior (narrative or meta-analytic) reviews in this area, search several electronic databases in which you believe relevant studies might exist (ensuring that these electronic searches include searches of unpublished studies such as dissertations), several listings of unpublished studies (i.e., conference programs, funding databases, and any research registries that exist in your field), and send out a request to authors via e-mail or listserv/website postings.

Second, you should create a list of studies obtained from these sources and ask some colleagues familiar with this research area to examine this list along with your inclusion/exclusion criteria. If they view it as complete, you have a good beginning. However, if they identify studies that are missing but should have been found, you should revise your search strategies (e.g., specifying different key words for electronic searches) and repeat the prior step.

The third suggestion is to take this list and begin forward and backward searches. You might start with forward searches, as this is less time-consuming. Here, you would start with a small number of the most seminal works in the area (in the absence of a clear idea of the seminal works, you might create a short list of the first studies and the studies published in the

top journals in your field). After performing forward searches with these seminal works (spending considerable time reviewing the citing papers to ensure relevance, as these types of searches are usually low in precision), you probably will have identified some additional studies; if not, you can reasonably conclude that forward searching will not yield any additional studies. Then, you can begin performing forward searches with the remaining studies, perhaps starting with the oldest studies first, as these have existed for the longest time and have therefore had more opportunity to be cited. At some point, you will likely reach a point where forward searches of more articles no longer yield new articles, and you can stop forward searching.

At this point, you can begin coding studies (see Chapters 4–7). While doing so, you should also perform backward searches (i.e., reading the works carefully for citations to other potentially relevant studies). My experience is that I often find a considerable number of additional studies when I begin coding, but that this number quickly diminishes as I progress in coding studies. If you find that you are almost never identifying additional studies near the end of your coding, you can be reasonably confident that your search is approaching exhaustion.

Despite this confidence, I recommend two additional steps to serve as a reality check. First, sit down with a few years of journals that are likely to publish studies relevant to your meta-analysis, and simply flip through the tables of contents and potentially relevant studies.[13] If you do not find any additional articles, then this adds to your confidence that you have conducted an exhaustive search. However, if you do find additional articles, then you obviously need to revise your search procedures (if you find relevant articles, carefully consider why they were not found—e.g., did the authors use different key words or terminology than you used in your search?). The second step, if your flipping through the journals suggests the adequacy of your search, is to send the list of studies again to some experts in your field (preferably some who did not evaluate the initial list). If they identify studies you have missed, you should revise your search procedures; but if they do not, you can feel reasonably confident that your search is adequate.

My intention is not to be prescriptive in the process you should take in searching the literature. In fact, I think that the search process I described is more intensive than that used for most published meta-analyses. Nevertheless, I present these steps as a model of a process that I believe leaves little uncertainty that your search is "close enough" to exhaustive. Although there is no guarantee that you have obtained every study from the population, I believe that after taking these steps you have reached a point where more efforts are unlikely to identify additional studies and are therefore not worth-

while. I also believe that no other potential meta-analyst would be willing to engage in significantly greater efforts, so your search represents the best that is likely to be contributed to the field. Moreover, by consulting with experts in your field, you have ensured that your peers view the search as reasonable, which usually means that reviewers will have a favorable view during the review process, and readers will view it as adequate after it is disseminated. In sum, I believe that strategies similar to the one I have described can provide a high degree of confidence that your search is adequate.

3.5 PRACTICAL MATTERS: BEGINNING A META-ANALYTIC DATABASE

Aside from perhaps persistence and patience, the most import virtue you can have for searching the literature for a meta-analysis is organization. As you have likely inferred, searching for studies is a time-intensive process, and you certainly do not want to add to this time by repeating work because of poor organization.

A good organizational scheme for the literature search will include several key components. First, you should have a clear, written statement of the inclusion/exclusion criteria that you will use in evaluating studies found through this search. Toward this end, it might be useful to record studies identified in your search that were excluded for one reason or another (recording why they were excluded). Second, you should have a clear list of methods for searching the literature, with enough details to replicate these searches. For example, you might have a list that begins:

Step 1: Read the following review papers and chapters (listing these works).

Step 2: Search the PsycINFO database using the following key words (listing the key words, including any wildcard marks and logical operations).

Step 3: Search the ERIC database using the following key words (listing the same set of key words as the step 2 search, unless there is reason to use other key words or logical operations).

You then record the dates—and names, if multiple people are conducting the searches—of each search.

During the course of these searches, you will scan many titles and abstracts in an attempt to determine whether each study is relevant for your

meta-analysis. I suggest that you be rather inclusive during this initial screening, retaining any studies that *might* meet your inclusion criterion. You should also retain any nonempirical works, such as reviews or theoretical papers; although these do not provide empirical results for your meta-analysis, it will be worthwhile to read them (1) to identify additional studies cited in these papers, and (2) to inform interpretation of results of your meta-analysis.

As you are identifying works that you will retain, it is critical to have some way of organizing this information. I use spreadsheets such as that shown in Table 3.1. (I have shown only four studies here, your spreadsheet will likely be much larger.) Although you should develop an approach that meets your own needs, this example spreadsheet contains several pieces of information that I recommend recording. The first column contains a number for each paper (article, chapter, dissertation, etc.) identified in the search. The number is arbitrary, but it is useful for filing purposes (as the number of papers becomes large, it is useful to file them by number rather than, e.g., author name). The next four columns contain citation information for the paper. This information is useful not only for citing the paper in your write-up, but in identifying repetitive papers during your multiple search strategies (for this purpose, having this information in a searchable spreadsheet is useful). The sixth column contains the abstract, which is useful if you want to search for specific terms within your spreadsheet. I recommend copying this information into your spreadsheet if it is electronically available, but it probably is not worth the time needed to type this in manually. The seventh column identifies where and when the paper was found; recording the date is important because (1) you might want to update the search near the completion of your meta-analysis, and (2) you should report the last search dates in your presentation of your meta-analysis. The two rightmost columns (columns eight and nine) contain information for retrieving and coding the reports. One column indicates whether you have the report, or the status of your attempt to retrieve it (e.g., the third paper notes that I had requested this dissertation through my university's interlibrary loan system). The last column will become relevant when you begin coding the studies (see Chapters 4–8). Here, I have recorded the person (BS = Brian Stucky, the second author on this paper) who coded this report and the date it was coded. Recording both pieces of information are valuable in case you later identify a problem in the coding (e.g., if one coder was making a consistent error) or if you revise the coding protocol (you then need to modify the coding of all studies coded before this change). In this column, I also record when studies are excluded for a particular reason; for instance, the fourth study was excluded because it used an adult sample (which was one of the specified exclusion criteria in this review).

TABLE 3.1. Example Spreadsheet for Organizing a Literature Search

Paper No.	Authors	Year	Title	Source	Abstract	Found in	Have	Entered
1	Crick and Grotpeter	1995	Relational aggression …	*Child Development, 66*(3), 710–722	Assessed a form …	PsycINFO (Nov. 2005)	Yes	BS, 12/1/05
2	Hawley, Little, and Card	2007	The allure of a mean friend …	*International Journal of Behavioral Development, 31*(2), 170–180	Recent theory …	PsycINFO (May 2007)	Yes	NC, 9/12/07
3	Blachman	2003	Predictors of peer …	Dissertation, University of California, Berkeley	Examined the role …	Proquest dissertation	Requested ILL	
4	Bailey and Ostrov	In press	Differentiating forms …	*Journal of Youth and Adolescence*	The purpose …	E-mail request	Yes	Adults

Note. The table lists Bailey and Ostrov as "in press" even though it was published in 2008. I left the date in the table as "in press," however, because the table is meant to show progress as it occurred during the time of this research (which was prior to this work being published).

Of course, you may use a different way of organizing information from your literature search. The point is that you should have *some* way of organizing information that clearly records important information and avoids any duplication of effort.

3.6 SUMMARY

One of the most important steps of a meta-analytic review is obtaining the sample of studies that will provide the data for your analyses. To define this sample, we need to specify a clear set of inclusion and exclusion criteria specifying what types of studies will and will not comprise this sample. We then search the literature for studies fitting these inclusion criteria. Several approaches to searching for literature exist, and I have described some of the more common methods. The goal of this search is to obtain an unbiased, typically exhaustive (i.e., complete) sample of studies.

3.7 RECOMMENDED READINGS

Reed, J. G., & Baxter, P. M. (2009). Using reference databases. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 73–101). New York: Russell Sage Foundation.—This chapter provides a very detailed, practical guide to using electronic databases, including forward search databases.

Hopewell, S., Clarke, M., & Mallett, S. (2005). Grey literature and systematic reviews. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 49–72). Hoboken, NJ: Wiley.—This chapter describes several ways of identifying and retrieving studies that are more obscure than traditional journal articles, and discusses the biases potentially introduced by not including this literature.

NOTES

1. The details (e.g., effect sizes, distributions around the mean) of this example will become clearer as you read subsequent chapters. For now, you should just try to understand the gist of this example.
2. In principle, a meta-analysis does not need to include all studies that exist. Instead, you can select a random sample of all existing studies on which to perform your analyses, assuming the studies you have selected provide adequate

statistical power to evaluate your research questions. I view this type of random sampling as an extremely valuable approach to performing reviews in areas where there is so much empirical literature that a full meta-analysis is not practical. However, very few meta-analytic reviews use this random-sampling approach; nearly all attempt to be exhaustive in their inclusion of studies. Unfortunately, this typical practice of being exhaustive seems to have created a standard where meta-analytic reviews are expected to be exhaustive, and the random-sampling approach would likely draw criticism.

3. The importance of developing clear operational definitions of constructs is important regardless of effect sizes used, whether they are of single variables (e.g., means or proportions) or multivariate effect sizes (see Chapter 7).
4. If you are particularly interested in drawing cross-cultural conclusions and there exists adequate numbers of studies written in a tractable number of languages, it may be possible to hire translators. However, you should remember that coding studies is an intensive effort (see Chapters 4 and 5) that requires considerable technical expertise. Because it would be difficult to find someone with both multilingual and meta-analytic skills, and require considerable amounts of their time, this is not a viable alternative in the vast majority of cases. For this reason, restriction of populations of studies to those written in languages you know is often reasonable as long as you recognize this restriction.
5. This condition is necessary to include a study in your analyses. However, you should also consider whether the studies that report insufficient information differ in meaningful ways, with the most relevant possibility being that the results were nonsignificant. If you find that a considerable number of studies report insufficient information to compute effect sizes (and other efforts, such as contacting the authors, do not alleviate this problem), then you should report these studies in your report for transparency.
6. Here, performing the meta-analysis with a random sample of studies might be preferable to changing your inclusion/exclusion criteria, especially if doing so makes the population of studies of lesser interest. Footnote 2 of this chapter describes some of the challenges to this approach.
7. To illustrate this cost, consider my experience when publishing the example meta-analysis I use throughout this book: During this review process, one of the reviewers suggested that I "plow through" the approximately 30,000 studies that could be identified using a very general search term like "aggression." Assuming 10 minutes to review each study for possible inclusion (which is a conservative estimate), this process would have taken over two years of 40 hours/week reviewing. During this time, approximately 3,000 additional studies identified using this search term would have been added, thus requiring another 3 to 4 months of full-time reviewing. Furthermore, during the coding, analysis, and write-up of these results, a couple thousand more works would likely have been

added to the database. Although this reviewer was certainly trying to be helpful by ensuring high recall, this example illustrates that the cost of low precision can be substantial in making a meta-analysis impossible.

8. The use of nonacademic search engines (e.g., Google scholar) might be especially plagued by inconsistency in what works are included. I personally do not use these nonacademic search engines. If you do decide to use one, I recommend *not* using it as a primary search method, but rather as a check of the adequacy of your other search procedures (i.e., after searching for literature using other methods, does this nonacademic search engine uncover additional works that should have been included?).
9. We did not do so in the actual meta-analysis because the number of studies using samples outside of this age range was reasonably small.
10. To my knowledge, no one has evaluated this possibility empirically. I also suspect that factors unrelated to the effect sizes (e.g., length of time since the presentation, your persuasiveness and persistence in requesting presentations) are more influential with regard to response than the effect sizes. But this possibility of biased response should be kept in mind when response rates are low, and it might be worthwhile to evaluate this possibility (through, e.g., funnel plots or effect size–sample size correlations; see Chapter 11) among the conference presentation included in your meta-analysis.
11. I do not believe that anyone has evaluated this empirically.
12. I find it comforting to consider that, just as there has never been a flawless study (see quote by Cooper, 2003, in Chapter 2 of the present volume), there has never been—and never will be—a flawless meta-analysis. Although you might strive to obtain every study within your sample, there comes a point of diminishing returns where a tremendous amount of additional effort yields very few additional benefits. When this point is reached, your field benefits more from timely completion and dissemination of your meta-analysis than futile efforts to obtain additional studies.
13. This image might seem quaint to some readers. If you prefer, point-and-click your way through the online tables of contents of some relevant journals.

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com