

2008

Cumulating Evidence about the Social Animal: Meta-Analysis in Social-Personality Psychology


Blair T. Johnson Dr.

University of Connecticut - Storrs, blair.t.johnson@uconn.edu

Marcella H. Boynton Dr.

Duke University, mhb23@duke.edu

Follow this and additional works at: https://opencommons.uconn.edu/chip_docs

 Part of the [Family, Life Course, and Society Commons](#), [Health Communication Commons](#), [International and Intercultural Communication Commons](#), [Interpersonal and Small Group Communication Commons](#), [Mass Communication Commons](#), [Nature and Society Relations Commons](#), [Politics and Social Change Commons](#), [Psychology Commons](#), [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#), [Race and Ethnicity Commons](#), [Social Influence and Political Communication Commons](#), [Social Psychology and Interaction Commons](#), [Sociology of Culture Commons](#), and the [Theory, Knowledge and Science Commons](#)

Recommended Citation

Johnson, Blair T. Dr. and Boynton, Marcella H. Dr., "Cumulating Evidence about the Social Animal: Meta-Analysis in Social-Personality Psychology" (2008). *CHIP Documents*. 31.
https://opencommons.uconn.edu/chip_docs/31

A corrected version of this paper was published (and should be cited as):

Johnson, B. T., & Boynton, M. H. (2008). Cumulating evidence about the social animal: Meta-analysis in social-personality psychology. *Social and Personality Psychology Compass*, 2, 817 -841. (doi: 10.1111/j.1751-9004.2007.00048.x)

Cumulating Evidence about the Social Animal: Meta-Analysis in Social-Personality Psychology

Blair T. Johnson* and Marcella H. Boynton
University of Connecticut

Abstract

Like most scientific fields, social-personality psychology has experienced its own explosion of research related to such central topics as aggression, attraction, gender, group processes, motivation, personality, and persuasion, to name a few. The proliferation of research can be a monster unless it is tamed with the scientific review strategy of meta-analysis, ~~which are~~ literally analyses of past analyses that produce a quantitative and empirical history of research on a particular phenomenon. The purpose of this paper is to outline the basic process and statistics of meta-analysis, ~~particularly as it~~ pertains to social-personality psychology. ~~The many detailed steps of meta-analysis can be reduced to four:~~ (i) defining the problem under review; (ii) gathering qualified reports and putting their findings and methods into a database, (iii) analyzing that database, and (iv) interpreting the results and reporting them. Use of meta-analytic strategies has paralleled the knowledge explosion in social-personality psychology, but must be used and consumed with careful discernment if the cumulated evidence about the social animal, *Homo sapiens*, is to have maximal value.

At the focus of the field of social-personality psychology is the consummate social animal, *Homo sapiens*. Like any scientific field, progress in social and personality psychology hinges on careful accumulation of knowledge via series of studies that evaluate hypotheses addressing how and why the social animal behaves as it does. For more than 100 years, personality and social psychologists have conducted myriad research studies examining human behavior, ranging from aggression to attraction, from gender to group processes, and from personality to persuasion (Richard, Bond, & Stokes-Zoota, 2003). In the process, the discipline has expanded to a field interested in any and all psychological phenomena involved in intrapersonal ~~or~~ interpersonal processes. Reflecting the explosive growth of the field, the number of peer-reviewed journals specializing in matters of social-personality psychology has increased from a mere handful in the 1930s and 1940s to more than 40 today (Journal Citation Reports, 2007).

With so much research now available in personality and social psychology, it is a seemingly overwhelming task to stay abreast of the most current

research developments in the field. Moreover, the results of different studies focusing on a given topic often appear to conflict and often really do conflict. How then does a researcher make sense of the literature? As we review in this article, the answer to this question is increasingly to employ *meta-analysis*, which is literally an analysis of past studies' analyses and is also sometimes termed a *research synthesis* or a *quantitative review*. These broader terms encompass not only the statistical but also the surrounding steps that constitute a meta-analytic review. Although meta-analysis is becoming an essential tool in the researcher's methodological 'tool-box', many researchers remain relatively naïve to these procedures and statistics. In this paper, we first document the knowledge explosion in social-personality psychology, then introduce practices and processes of meta-analysis in social-personality psychology (and related fields), and conclude with some recommendations for future reviews related to the social animal. In summary, we strive to provide a starting point for those who wish to conduct their first meta-analysis as well as an aid for improved comprehension of meta-analytic reviews for all readers.

The Knowledge Explosion in Social-Personality Psychology

The knowledge explosion that characterizes nearly all scientific disciplines also typifies the field of social psychology, as even a cursory survey shows. Some current social-psychological examples include a review meta-analyzing 126 studies correlating implicit and explicit measures (Hofman, Gawronski, Gschwendner, Le, & Schmitt, 2005), a second meta-analyzing 515 studies (with 713 independent samples) examining intergroup contact theory, a third meta-analyzing 135 studies (with 163 independent samples) associating traits from the 5-factor model of personality with overall job satisfaction (Judge, Heller, & Mount, 2002), and a fourth meta-analyzing 92 studies (with 113 independent samples) tracking mean levels of change in personality traits across the lifespan (Roberts, Walton, & Viechtbauer, 2006). This small sampling of meta-analyses is indicative of the avalanche of recent published research on a plethora of social-psychological topics that currently inundate journals, book chapters, and other publication venues.

Richard et al.'s (2003) review of meta-analyses in social psychology provides an even larger sense of the knowledge base that personality and social psychology scholars have accumulated in recent times. They found more than 322 meta-analyses containing more than 25,000 studies utilizing more than 8 million research participants! Note that despite the documented existence of over 300 meta-analyses, there are numerous topics that have yet to receive meta-analytic attention. As a consequence, new primary-level studies appear at an ever-increasing rate of growth. Searches of PsycINFO also reveal the enormous empirical growth of the field. As two examples, there are more than 135 times more reports on the subject

of 'attitude' today than there were in 1960, ~~with more than~~ 19,500 entries. With more than 14,500 entries, there are more than 1200 times more reports on the subject of attribution. In both cases, the synonyms of these terms were not considered, reflecting the enormous bodies of literature that speak to these two example topics.

When one considers the sheer numbers of relevant reports coupled with the reality that most of the articles in leading journals increasingly contain multiple studies and that often studies' results seem to vary substantially, it becomes readily evident that to review the evidence related to any facet of the social animal can be a daunting task—indeed. Indeed, unless our research evidence is tamed using scientific review procedures, many evidential domains can amount to unwieldy leviathans. If social and personality psychology findings are not meaningfully cumulated, then the field runs the risk of inefficient knowledge gathering, the result being that we will increasingly reinvent the wheel. Consequently, meta-analysis has growing importance to the field's understanding and contextualization of past research findings as well as the identification of key factors that relate to the magnitude of differences between studies. Furthermore, given the burgeoning numbers of available meta-analyses, even if one does not wish to conduct a meta-analysis him- or herself, it is still increasingly vital to learn the skills necessary to be effective consumers of these reviews.

Meta-Analytic Methods

Traditionally, researchers have relied on the qualitative approach of narrative reviewing to synthesize literature on a given topic. Using this method, reviewers construct an integrated narrative summary of the results of individual primary studies, which might include a count of the number of studies that had either produced or failed to produce statistically significant findings. Qualitative reviews certainly have an important role to fill in providing general topical overviews (e.g., textbook summaries, introductions to journal articles reporting primary research) and assisting in theory development, for which the goal is less to be comprehensive than it is to show how variables *may* relate in a certain fashion for the phenomenon in question (Baumeister & Leary, 1997; Cook & Leviton, 1980). Nonetheless, narrative reviews conducted to reach conclusions about the nature of a particular phenomenon are typically limited in that they gather only a convenience sample of studies, lack statements about which study characteristics were considered or about how the quality of the studies' methods was evaluated, have difficulty reaching clear conclusions about what differences in study methods best explain disparate findings, and are overly reliant on statistical significance, minimizing the importance of the effect size magnitude. Consequently, narrative reviewers often reach erroneous conclusions about a phenomenon.

To demonstrate some of the limitations of qualitative reviews, Cooper and Rosenthal (1980) randomly assigned 32 social-personality graduate students and 9 faculty members to be narrative reviewers or meta-analysts for a small literature of 7 studies concerning a gender difference, asking them whether the independent variable in the study significantly related to the dependent variable. Narrative reviewers usually failed to reject the null hypothesis, but all of the meta-analysts concluded that a significant difference existed. If narrative reviewers can make mistakes even when the number of studies is small, one can imagine what mistakes occur for large literatures. By showing how researchers have addressed the phenomenon and delineating how results may have changed over time, meta-analysis has emerged as a leading tool to examine empirical histories of a phenomenon (e.g., Marcus-Newhall, Pedersen, Carlson, & Miller, 2000).

Although other sources might break the methods of meta-analysis into more detailed steps (e.g., Cooper, 1998; Johnson & Eagly, 2000; Lipsey & Wilson, 2001), for simplicity, we describe meta-analysis with four: (i) defining the problem and setting the boundaries of the theoretical domain, (ii) obtaining the relevant studies and coding them for their methods and effect sizes, (iii) analyzing the resulting meta-analytic database, and (iv) interpreting and reporting the results. In many ways these steps mirror the four parts outlined in a traditional research report: (i) introduction, (ii) method, (iii) results, and (iv) discussion. Figure 1 compares and contrasts the steps involved in implementing a meta-analysis to sections of a research report, which illuminate issues that a researcher might consider when writing or reading a meta-analysis, and considering audiences that vary in amounts and kinds of expertise.

As with primary-level research, the success of each step depends on those that precede them, and early steps should be implemented with an eye for the later ones. For instance, it is easier to locate relevant studies if the analyst has first done a good job of defining the meta-analytic question and reviewing relevant theoretical domains. Failure to keep the big picture in mind can easily result in difficulties achieving the specific steps. For example, defining a problem too broadly typically leads to a research literature that covers too many disparate areas to be quantitatively integrated in a meaningful or succinct way, and thus is sure to take longer to complete than is desirable. In such instances, and as a general rule, it is better to have collaborators than to proceed alone, dividing the work to create greater efficiencies and to create checks-and-balances for the work performed (e.g., so that reliabilities in coding may be calculated and presented). Given the complex decisions and judgments that meta-analysts usually face, the 'buddy system' has a great deal to recommend it. That said, analysts who decide to conduct a meta-analysis without a review team should be certain to select a project that is manageable in size and scope.

Consideration	Section of meta-analytic report			
	Introduction (Step 1)	Method (Step 2)	Results (Step 3)	Discussion (Step 4)
1. Objectives	Define the research phenomenon in question and conceptualize factors that relate to variance in that phenomenon. Make predictions about the phenomenon.	Retrieve all studies that match the criteria of the review, code them on meaningful dimensions, put study results on a single common metric (if necessary).	Discover meaningful patterns in study outcomes. Test the hypotheses, do exploratory analyses as necessary.	To interpret the findings in relation (a) to the goals of the review and (b) to previous knowledge of the research phenomenon.
2. Key methods	Reading of key past literature; setting goals for the review that determine the methods to be followed and the study dimensions to be considered.	Search for literature using available digital and analog strategies. Systematically code features of each study and its setting. Select ES type and calculate for each study.	Select analyses to model the ESs and to detect potential outliers among them. Use graphical displays as relevant.	Estimate the new contribution of the results to knowledge about the phenomenon. Weigh strengths and weaknesses of the evidence and the methods used to synthesize the information.
3. Any reader might ask:	Should I care about this problem? What problem are they considering, anyway? Does it have interesting, well-developed hypotheses? Can this review address my unique concerns?	How comprehensive was the search? Are interesting study dimensions coded? Do I understand the ES?	What is the overall trend of the phenomena? Do the outcomes vary more than expected by sampling error? Are the hypotheses supported?	Does it tell an interesting and convincing story? Does it provide novel insights into the phenomenon?
4. Other experts might ask:	Is the review too broad? Too narrow? Is this review relevant to the research I performed or maybe planning? Have the analysts reached interesting and testable predictions? Can this review address my unique research concerns?	Could search strategies have been performed better? Are studies selected fairly? Are there important biases in the methods? Is the ES chosen the best to examine study patterns in relation to the meta-analysis's goals?	Do the studies seem like appropriate exemplars of this phenomenon? Do outliers influence the results? Are the predictions tested in the best possible way? Do analyses eliminate potential confounds as explanations of the results?	Do I agree with the review's conclusions? Do biases in the methods change interpretations or leave unanswered questions? Are there important future directions that appear valuable to pursue?
5. Meta-analytic experts might ask:	Are the variables in question well-defined?	Is the sample of studies representative or exhaustive? Is it possible to integrate all studies? Is the ES type well-matched to each study's design?	Are ESs independent? Are meta-analytic models well-specified? Is variability in ESs modeled? Are all moderator variables analyzed? If a sufficient number of studies (k) are present, have moderators been modeled simultaneously?	Do the conclusions of the meta-analysis match its results? Are statistical and methodological limitations clearly stated?

Figure 1 A guide to the steps in meta-analysis considering the key considerations typically addressed in sections of a meta-analytic report (ES = effect size; k = number of studies).

1 *Defining the problem and determining the boundaries of the theoretical domain*

2 Selecting and defining the research domain to be analyzed is in some sense
3 the most complex part of meta-analysis. This choice sets the stage for all that
4 is to follow and is defined by both theoretical and practical considerations.
5 Essential to this conceptual analysis is a careful examination of the history of
6 the research problem and of typical studies in the literature. Being able to
7 have such a comprehensive understanding of the literature under consideration
8 assumes more than a passing familiarity with it. As a consequence, at least
9 some form of review ~~often needs to~~ be completed before the meta-analysis
10 can commence. This initial step is particularly important for social-personality
11 domains because scholars in these fields strongly embrace theory development
12 and testing (see Fiske, 2004). For a review to be published in one of the
13 key journals of social-personality psychology, it must embed its predictions
14 in predominant theories and show sophistication about those theories.

15 During this first step in the process, an analyst team should have a good
16 sense both of the debates and unanswered questions that surround the
17 phenomenon and of the extent to which meta-analytic procedures could
18 address them. Careful reading of past reviews (whether meta-analytic or narra-
19 tive) can be instructive, as can an inspection of oft-cited articles reporting
20 on primary-level studies. As an example, Durantini, Albarracín, Mitchell,
21 Earl, and Gillette (2006) wished to understand how social influence agents'
22 identities relate to their targets' behavior change. As opposed to considering
23 *all* behavior change contexts, these scholars focused on one of great applied
24 interest, HIV prevention. Additionally, rather than focusing on all of the
25 many available prevention measures, they ~~specifically focused on~~ condom
26 use. Having conducted previous reviews on the subject (e.g., Albarracín et al.,
27 2005), Durantini and her colleagues recognized that expert communicators
28 often differ from their target audiences in ways that ought to undermine
29 behavior change. For example, experts are often much more educated than
30 their targets and of higher socioeconomic status. Because the ~~list~~ of relevant
31 studies was large and offered many different combinations of communicator
32 and target audiences, it offered good potential to test predictions about
33 when behavior change should be large and when it should be small. Thus,
34 Durantini and her colleagues embedded their hypotheses about behavior
35 change in theoretical considerations stemming from several fields, including
36 social psychology. Using this strategy maximizes theoretical impact and
37 minimizes reliance on post hoc analysis strategies.

40 *Find relevant studies and code their methods and outcomes*

41 *Finding relevant studies.* Clearly, only some studies will be relevant to the
42 conceptual relation that is the focus of a meta-analysis, so analysts must be
43 able to define appropriate and meaningful inclusion criteria for the sample
44 of studies, which follow from the first step: defining the problem. These

1 criteria define the boundaries of the literature in a way that specifically
2 spell out what studies, or portions of studies within the reports, should
3 be included or excluded. To illustrate, Durantini et al. (2006) selected
4 reports that (i) included a condom-use-promotion intervention, (ii)
5 provided information about the interventionist, (iii) measured condom
6 use after the intervention, and (iv) offered appropriate statistics to calculate
7 effect sizes. This last criterion is so axiomatic to the process of meta-
8 analysis that it is often goes unstated. Developing selection criteria is often
9 a continuing process that evolves as the ~~meta-analyst examines~~ more studies;
10 ~~thereby discovering~~ the full range of research designs employed to
11 investigate a particular hypothesis.

12 A meta-analysis must also take study quality into account at an early
13 point in order to determine the typical ways a construct is operationalized
14 as well as ~~that~~ variables and results tend to be reported across a sampling
15 of the studies. Indeed, as we alluded, study quality is often a study selection
16 criterion in and of itself. Having carefully considered the literature, analysts
17 begin to gather studies with an eye toward knowing which studies may
18 have higher quality than others. Historically, scholars have commonly
19 assumed that only peer-reviewed and published articles have satisfactory
20 quality. Yet, as Begg (1994) reviewed, there are many reasons why a study
21 may be published, and only one of them is its quality. Many high-quality
22 studies are conducted but ~~are~~ never submitted for publication, including
23 many dissertations. Similarly, many published studies lack high quality,
24 even some featured in prominent journals. Because including more studies
25 generally increases the value of a meta-analysis, it is advisable to try to
26 locate *all* the studies that might be suitable for inclusion, whether or not
27 they have been published. Typically, published studies report larger rela-
28 tions than unpublished studies (Lipsey & Wilson, 1993); thus, including
29 unpublished studies usually eliminates some bias and also gives the analyst
30 a direct way to examine for publication bias – by comparing unpublished
31 studies to published studies.

32 Meta-analytic convention is inconsistent about whether to define
33 selection criteria so as to exclude studies of low quality or to include them
34 and examine whether the flaws relate to the phenomenon under investi-
35 gation (Saunders, Soomro, Buckingham, Jamtvedt, & Raina, 2003). Both
36 strategies have advantages and disadvantages. Restricting a sample to those
37 studies judged to have high quality may have the advantages of less
38 database work (due to a smaller sample) and findings that are easier to
39 explain (because presumably methodological 'noise' is removed), yet meta-
40 analyses based purely on 'high-quality' studies may be more open to the
41 criticism that they are too selective or reflect the analysts' idiosyncratic
42 views of what constitutes a high-quality study. The alternative strategy,
43 including all studies, reverses these advantages, and disadvantages. Now
44 it is imperative on the analysts to show the extent to which the meth-
45 odological features that may tap study quality actually relate to the

phenomenon by using these dimensions as moderator variable in the later analyses. Moreover, such reviews may well involve more work and result in a set of findings that is more difficult to explain. The choice of strategy is up to the analysts: If a methodological review is a priority for the literature in question, then the broader strategy is indicated; otherwise, the narrower strategy is probably more advisable.

Common methods for finding studies include examining the reference lists of existing reviews, consulting general journal search engines such as PsycINFO, PubMed, Web of Science, ERIC, and Dissertation Abstracts, obtaining published sources that have cited seminal articles within the literature, contacting the extant network of researchers who work on a given topic to ask for new studies or unpublished studies, posting advertisements for unpublished manuscripts and datasets on professional listservs, and by manually searching important journals to find some reports overlooked by other search techniques (see Johnson, Scott-Sheldon, Snyder, Noar, & Huedo-Medina, *forthcoming*). Additionally, Internet search engines such as Google are increasingly valuable in literature searches because they can help find reports that do not appear in the other databases (especially unpublished reports).

Researchers are inventive and innovative when testing their ideas. The good news is that this diversity fulfills the goal of maximizing the use of multitrait, multimethod research, which offers optimal discriminant and predictive validity (Campbell & Fiske, 1959). In other words, a diversity of studies can also often address the external validity of the findings. The bad news is that, at its worst, the inclusion of a wide range of methods can make a meta-analysis too large and complex to carry out in a reasonable time frame; it is also likely to increase the extent to which study findings conflict. In such instances, reviewers may divide a literature into two or more research syntheses, each addressing a different aspect of a broad research question. Alternatively, a meta-analyst might obtain more resources to increase the feasibility of a particular meta-analysis (e.g., giving academic credits to students to assist, obtaining grants for hiring personnel).

Despite the potential for conflicting findings across studies, meta-analyses typically do incorporate studies with differing methods. For example, researchers may use self-reports, close others' reports, and objective outcomes to assess a particular behavior. Although some would contend that objective outcomes (e.g., actual weight loss) are the 'best' operationalization of behavior changes, it is not usually advisable to omit the typically more copious subjective measures (e.g., perceived weight loss). Analysts should make this decision at the outset of their meta-analysis, considering whether restricting an analysis to a certain sort of measure would artificially narrow the potential results of the review. It is of course possible to include multiple operationalizations as long as they all pertain to the original research question. When meta-analysts take this route, they

should make sure to code for the type of research method in question so that it can be included in later analyses (e.g., Anderson, 2004).

For some research questions, it may be appropriate to select studies using geographic setting, culture, study population, and other factors. The questions some meta-analyses pursue make having diversity in these dimensions a virtue, whereas the questions other meta-analyses pursue can make them a liability. An analyst interested in, say, how cultural factors underlie social influence wants to maximize variability on those dimensions such as a cross-cultural meta-analysis of studies using Asch's line-judgment conformity task did (Bond & Smith, 1996). In contrast, an analyst concerned with reviewing how a particular population in a specific locale reacts to a question or paradigm will want to eliminate populations and locales that lie outside that area, as Twenge, Zhang, and Im (2004) did in their meta-analysis of US studies on locus of control (the extent to which individuals believe controlled by outside forces rather than their own efforts).

Another issue associated with the defining of inclusion criteria is whether to include reports written in alternative languages, which often are defined as non-English studies. ~~Non-English studies~~ might usually comprise a very small minority of research reports for the topic under consideration, in which case it is unlikely that their exclusion would change meta-analytic results. ~~Note that exclusion of~~ reports from other languages has the potential to introduce ethnocentric bias, which should be avoided as much as possible. Hence, whether or not non-English reports are in the minority or majority, every effort should be made to include them if they otherwise match the selection criteria. Translation services can be obtained at minimal or no cost by interacting with individuals or networks that have the requisite skills (e.g., students from the country in question, language departments); ~~similarly,~~ software to perform translation often can do an adequate job relaying the essential information that is required for the analyses.

Whatever selection criteria are used, a review team should carefully record their methods of locating studies, including such details as the names and databases ~~what~~ were searched, the time period covered for each database, and the keywords/search terms used. Details of the search procedure should be included in the methods section of the meta-analysis report, to enable readers to make adequate judgments about the adequacy of the procedures used and to permit other analysts to replicate the search. Indeed, we recommend keeping an ongoing journal to record the many decisions made as the endeavor proceeds. These details gradually form the content of the Method section that will eventually ~~be the product of~~ the meta-analytic endeavor (see Figure 1).

Coding important features of the studies. A coding sheet or an electronic database needs to be created, pretested, and revised to capture important features of the studies and their contexts (see the Appendix in Lipsey &

Wilson, 2001, for an example coding form). All variables to be coded should be operationalized in a very specific manner. Study characteristics may consist of continuous variables with values existing along ratio, interval, or ordinal scales, or they can be categorical variables having discrete values reflective of meaningful qualitative differences. Just as in primary-level research, it is usually wise to avoid dichotomizing or otherwise 'dumbing down' the data at the coding stage. A master codebook can explain the details for each category and the information to be included in the database. As a best practice, two or more coders should code each study so that reliability can be assessed (e.g., Cohen's κ , Spearman-Brown r ; see Krippendorff, 1980) and included in the meta-analysis report.

~~It is often the case that~~ multiple independent samples reside in a single report, especially when the paper describes multiple 'studies'. ~~It is meta-analytic convention to~~ treat each substudy as an independent study. ~~It is often the case that~~ meta-analyses of social and personality psychological topics disaggregate the data from individual studies. A good example of reorienting data to address a meta-analyst's topic of interest is Johnson and Eagly's (1989) meta-analysis examining the influence of involvement on message-based persuasion. These researchers theorized that the size of this relation would depend on the strength of the persuasive argument. In order to test this idea, they created two 'studies' from each involvement study that manipulated argument quality: one study from the strong-arguments condition and another from the weak-arguments condition (they also divided studies on other bases). This strategy maximized the moderator analyses' ability to test the interaction between involvement and argument quality.

Potential *moderator* variables should be defined a priori and be kept to a minimum number; otherwise a meta-analyst can be accused of 'fishing' for significant results. A moderator variable is any variable that might potentially explain inconsistencies among the study findings, which we discuss in Step 3, below. These a priori variables can ~~pertain to~~ any characteristic relating to the study or its context, and they should flow from Step 1, the definition of the meta-analytic problem. Analysts may also code many other dimensions of studies for use as *descriptive* or *exploratory* variables. Reporting such information helps to describe the sample of studies and to place subsequently reported meta-analytic findings into context. These dimensions may also prove important to exploratory analyses performed after the planned analyses are complete.

Coded variables are limited only to the imagination of the analyst team: they might concern the population (e.g., *M* age, ethnic group proportions, *M* educational level, institutionalization status, etc.), the locale surrounding the study at the time it was conducted (e.g., racial climate, support for affirmative action), study methods (e.g., methods of recruiting participants, differences in measuring the dependent variable), or publication

dimensions (e.g., published/unpublished, language of publication). For example, to enable more exact comparisons among studies along the dimension of argument quality, Johnson and Eagly (1989) sampled from studies' arguments and asked undergraduate judges to assess this dimension. In this fashion, even studies that did not manipulate argument quality could enter into the analyses involving this dimension. A variation on this strategy can gauge study quality. For example, Shadish, Navarro, Matt, and Phillips (2000) examined the effects of psychological therapies on clinical populations. These scholars assessed clinical representativeness using 10 representativeness criteria, which they then used to code the populations and settings for each research study. Using this strategy, the researchers were able to conclude that effect sizes were not a direct function of clinical representativeness, which refuted conclusions from other research. Strategies such as these allow for the creation of ~~potential moderating~~ variables that ~~are not necessarily reported, or even considered,~~ in the individual studies.

Capturing study findings on a common metric (calculating effect sizes, if necessary). At the heart of the meta-analytic endeavor is capturing study findings on a metric that is common to each study. ~~Only one metric is chosen to capture~~ each study's findings, ~~although there are potential exceptions to this rule.~~ Social-personality ~~psychologists~~ usually report a variety of measures; in such cases, meta-analytic practice is to use a standardized effect size index to gauge the degree of association between two variables. This strategy loses the original units of the studies' scales, but the effect size still estimates the magnitude of the association for each study in a way that can be compared across studies. This procedure is especially advantageous given that statistical significance is usually not a good measure of the magnitude of the effect of interest (Kline, 2004). In other words, a nonsignificant result may in fact correspond to a large effect and a significant result to a small effect. In cases for which all studies used the same measure (e.g., mean levels of ~~social dominance orientation~~), meta-analyses may keep findings in their original scale units instead of converting them to a standardized effect size. As we ~~address~~ later in this subsection, this strategy offers information that the standardized effect size strategy cannot.

Although many readers assume that there are only one or two available effect size statistics, there are in fact a large array of potential effect sizes to gauge the linkage between two variables (~~see Lipsey & Wilson, 2001~~), as Figure 2 illustrates. For any particular pairing of two variables, there is at least one effect size form – if not many – that appropriately captures the linkage, which depends on the precision of each variable. The task of the analyst is to pick the effect size type that best captures variation in their phenomenon and for which appropriate weights can be calculated for the purposes of analyses (see Step 3). It is beyond the scope of this article to review all possible effect size types (see Lipsey & Wilson, 2001),

Nature of second variable	Nature of first variable		
	Continuous	Ordinal	Categorical
Continuous	<ul style="list-style-type: none">Pearson correlation (r)Standardized regression slopes (β)Unstandardized regression slopes	<ul style="list-style-type: none">Biserial correlation (r_b)	<ul style="list-style-type: none">Point biserial correlation (r_{pb})Standardized mean difference (d)Unstandardized mean effect size
Ordinal		<ul style="list-style-type: none">Spearman correlation (ρ or ρ_h)Tetrachoric correlation (r_{tc})	<ul style="list-style-type: none">Rank-biserial correlation (r_{rb})
Categorical			<ul style="list-style-type: none">Phi coefficient (ϕ)Odds ratio (OR)Risk ratio (RR)Risk difference

Figure 2 Potential two-variable effect sizes dependent on the intersection of the variables in question; the cells shaded in yellow indicate the most frequently examined associations in social-personality research.

but as yellow cells of Figure 2 highlight, social-personality psychologists typically deal with either two continuous variables, in which case the effect size to use is Pearson's r , or one categorical variable and one continuous variable, in which case the effect size to use is d . Thus, if an experimental group were compared to a control group, or men to women, d is the effect size of choice. If the key relation is instead two continuous scales, such as amount of interracial contact and amount of prejudice, a researcher would employ r . It is possible to use a member of the family of correlations in place of d , but if so, as the Figure 2 shows, it should be the point-biserial r and *not* the Pearson r (for further discussion, see Johnson et al., [forthcoming](#); Lipsey & Wilson, 2001; McGrath & Meyer, 2006). Whereas values of r can range from -1 to $+1$, values of d can take any negative or positive value. Signs are defined so that studies with opposite outcomes have opposing signs; effect sizes of 0 have no association (r) or difference (d).

Nearly all inferential statistics (e.g., t -tests, F -tests) and many descriptive statistics (e.g., means and standard deviations) can be converted into d or r or other effect size types (Johnson & Eagly, 2000; Lipsey & Wilson, 2001). Reports occasionally provide only inexact statistics relevant to the effect size (e.g., the difference was nonsignificant). Meta-analysts customarily contact the authors of the study to request the necessary information. In the absence of these data, the convention is to maintain the imprecise information so that the study is not lost to the review (Rosenthal, 1991). For example, an effect described as $P < 0.05$ can be interpreted as an exact P value ($P = 0.05$) to an effect size. This method has the advantage of keeping the report in the sample but it is likely to be a conservative estimate, especially when sample sizes are small.

Analyzing the meta-analytic database

Once all studies are coded and effect sizes calculated, the third step in the meta-analytic process is to synthesize the data. An exhaustive survey of general analytic approaches to meta-analysis is beyond the purview of this article, but further discussions and comparisons are available elsewhere (e.g., Field, 2001, 2005; Hall & Brannick, 2002; Hunter & Schmidt, 2004; Sánchez-Meca & Marín-Martínez, 1997). Here we focus on the general procedures that analysts conventionally use to analyze their data.

Overall tendencies – the mean effect size. As a first step to quantitative synthesis, all available study outcomes are combined by averaging the effect sizes, with each effect size weighted in some way, usually by the inverse of its variance (Hedges & Olkin, 1985). Weighting gives greater weight to the more reliably estimated study outcomes, which are in general those with the larger samples (e.g., Hedges, Cooper, & Bushman, 1992). If the effect size ~~share the same~~ effect size, then they would differ only by unsystematic sampling error. The test statistic Q evaluates this hypothesis using an approximate χ^2 distribution with $k - 1$ degrees of freedom, where k is the number of studies (Cochran, 1954). If Q is significant, the hypothesis of the homogeneity (or consistency) of the effect sizes is not accepted. In this event, the weighted mean effect size is an inappropriate statistic to represent the distribution of outcomes and it is likely that subgroups of effects exist; in other words, a significant Q value indicates that there are likely to be one or more significant moderating variables that explain differences in effect sizes. In this case, further analyses are warranted to test potential moderating variables responsible for the variability in effect sizes.

Q deserves careful interpretation, in conjunction with inspecting the values of the effect sizes. Even if the homogeneity test is nonsignificant, it is possible for significant moderators to emerge, especially when Q is relatively large (Johnson & Turco, 1992). If the number of studies is small, tests of homogeneity are known to have lowered power to detect the null hypothesis of homogeneity (Hardy & Thompson, 1998; Harwell, 1997). Although it is new to meta-analyses in social psychology, another gauge of consistency deserves consideration. Specifically, Higgins and Thompson (2002) introduced a homogeneity index, I^2 , based on Q and its degrees of freedom. Values of I^2 range from 0 to 100%, where high values indicate more variability among the effect sizes. Although I^2 is subject to the same conditions and qualifications as is Q (Huedo-Medina & Johnson, 2007; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006), its use eases ~~the way to compare~~ results of different meta-analyses or portions thereof.

An adjunct to homogeneity statistics is the use of graphical displays to examine the distribution of effect sizes. Funnel plots can help to diagnose

studies' effects that differ in magnitude and weight from the rest; these cases can be inspected to determine if they differ greatly from the others and can be removed if that is the case (Light & Pillemer, 1984). A funnel plot is a simple scatter plot with effect size on the X axis and its weight on the Y axis. It is termed a funnel because when the cases pattern with no bias, the shape resembles an inverted funnel. Of course, it is also possible that these outliers vary meaningfully on dimensions for which moderator analyses are planned. In this case, the practice is to maintain the outliers in the sample for further analyses. Inspection of the funnel pattern might also reveal gaps in the funnel pattern that theoretically should not be present: If the funnel lacks cases that would have nonsignificant effects, then one plausible explanation is that a publication bias exists (for statistical efforts to define these missing studies' effects, see Duval & Tweedie, 2000; Hedges & Vevea, 1996; Vevea & Hedges, 1995).

When conducting analyses of meta-analytic data, it is often the case that creating one or more figures can go a long way in succinctly summarizing the findings. For example, Figure 3 uses the basic data from a meta-analysis examining gender differences in social dominance orientation (Lee, Pratto, & Johnson, 2007), which concerns people's degree of support for social hierarchy and inequality. The figure shows six different ways that the very same study results may be depicted. First note that there is little visual evidence of any systematic underreporting of results because the findings appear normally distributed (panel a), including comparisons in the nonsignificant range (panel b). More interesting is the fact that the standardized gender difference results (panels a and b) were far more consistent than the unstandardized, arithmetic mean forms (see especially panels c and f); indeed Lee and colleagues found that homogeneity for the gender comparison was smaller ($I^2 = 65\%$) than that for the arithmetic means ($I^2 = 98\%$). Findings such as these reflect the fact that raw mean values are subject to any possible influence from the environment, whereas a standardized comparison between groups sampled from a common environment control for every influence except those related to the group comparison. In this illustration, the gender comparison reflects only the impact of gender and any other variable with which gender might be related. Panel f reveals that mean social dominance values for both males and females move across the entire range of the scale, yet the gender difference consistently favors males.¹

The impact of moderating variables. Almost all meta-analytic reviews in social-personality psychology seek to identify and explain differences in research findings by examining what coded variables (i.e., moderator variables) explain the differences. Specifically, a theory might predict that a third variable influences the relation between the independent and dependent variables. Using this third variable as a potential moderator of the effect, the analyst would code all of the studies for their status on

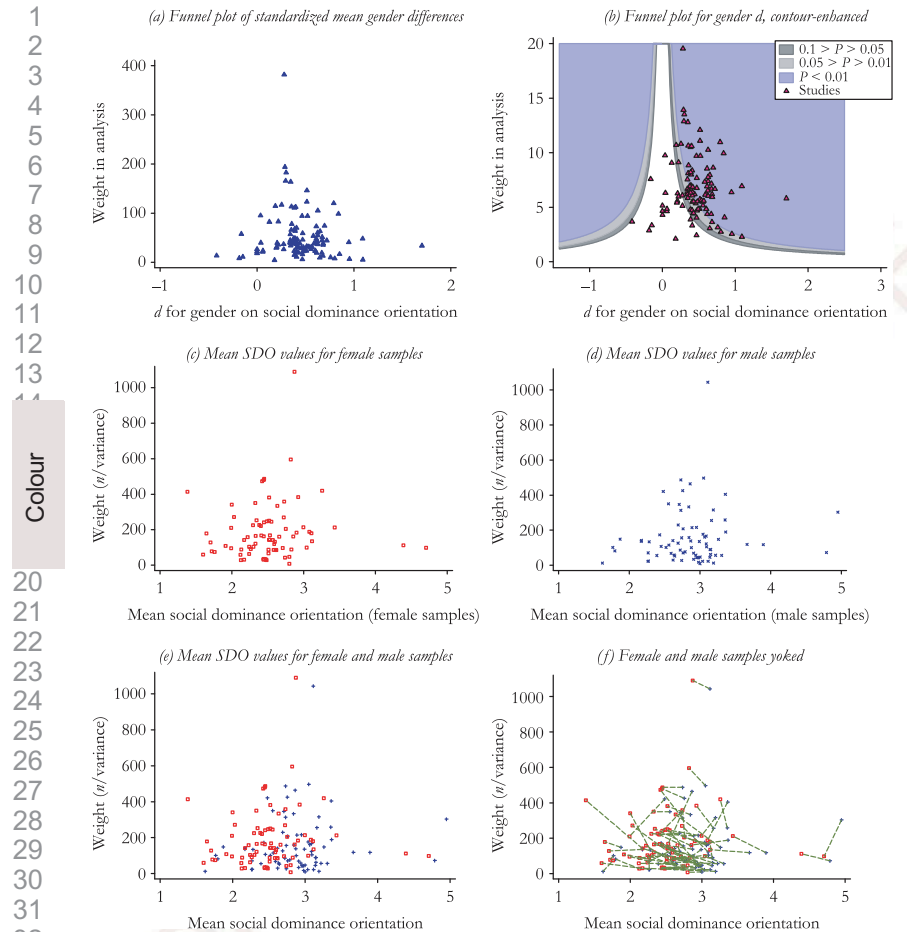


Figure 3 Social dominance orientation (SDO) from multiple studies gathered from around the world and offering a gender comparison (Lee et al., 2007); these graphs display the same data in various funnel plots: (a) gender differences in SDO displayed as a standardized mean effect size, d , where positive values imply that males exhibited more support for social dominance than females. (b) The same effect sizes with contours shown for regions of significance; note that none of the d s in the female direction attained significance. (c) Raw, unstandardized mean SDO values depicted separately for female and (d) for male samples; note that male means tend to be higher than female means. (e) The same values from panels c and d collapsed into one graph. (f) The same values from panels c through e, yoking each pair of means with lines; the fact that most lines have the male value to the right of the female value corresponds to the predominate pattern exhibited in panels a and b, but exhibits it in an unstandardized form. The slopes indicate whether males or females carry greater weight in the analysis (based on their sample size and observed variance).

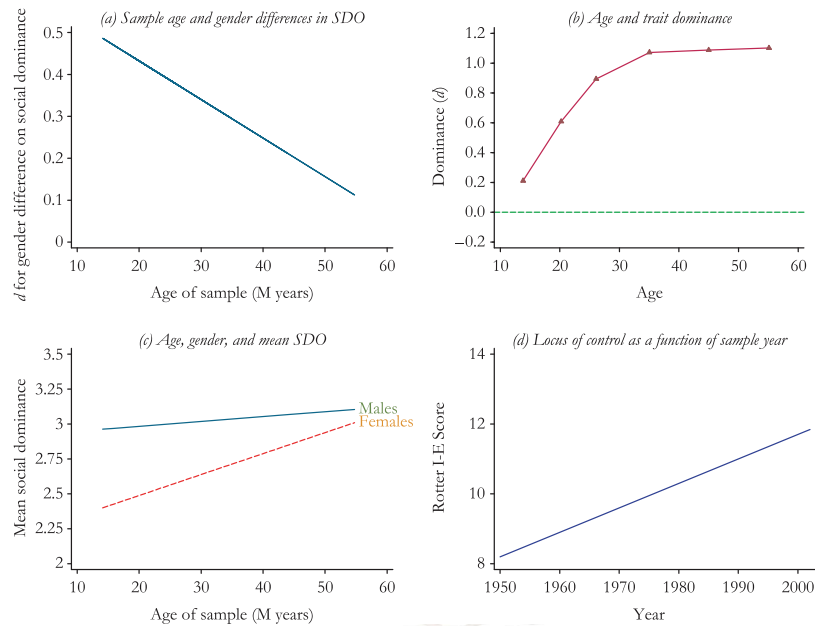


Figure 4 Selected meta-analytic findings displayed graphically. Panel a, gender differences in social dominance orientation narrow with increasing sample age (Lee et al., 2007). Panel b, dominance as a personality trait increases until middle age and then stabilizes (Roberts et al., 2006; the dotted green line shows the reference value at age 10). Panel c, both genders increase in social dominance across age of sample, but female samples increase more than male samples (Lee et al., 2007). (Note that this pattern helps to explain what is happening in the same data presented in standardized mean effect form in Panel a). Panel d, from 1950 to 2002, US samples expressed increasing externality as gauged by Rotter's I-E scale (Twenge et al., 2004).

the moderator. Moderator analyses can be conducted with categorical (analogous to analysis of variance), or continuous (analogous to regression analysis) variables; it is possible to create different predicted models using those categorical or continuous variables that explain greater amounts of variance. This *moderator variable approach* (or *effect modification approach*) examines whether the moderator relates to the study results in a systematic way. In so doing, analysts advance beyond the simple question of *whether* an independent variable relates to the dependent variable, to address the question of *when*, or under what circumstances, the magnitude or sign of the association varies.

One example of how moderator analyses can offer surprising and in-depth insight into a psychological construct are analyses conducted by Lee et al. (2007), who investigated and found that age of sample related to the gender difference in social dominance orientation, with the gender difference narrowing in older samples and at its widest for young samples (see Figure 4, panel a). Similarly, Roberts et al.'s (2006) meta-analysis of

personality change across the lifespan showed, among other things, that as people age, their social dominance scores increase, leveling out after about age 35 (panel b). As a final example, Twenge et al. (2004) showed how samples of young Americans have exhibited locus of control scores that are increasingly external across a 50-year span (Figure 4, panel d).

As we have implied, moderator analyses need not be limited merely to standardized effect size indexes. We have argued that examining unstandardized forms of study information can be valuable as well (see Figure 2). As a case in point, consider Lee et al.'s (2007) finding that as age of sample increased, gender differences decreased. Because of the nature of the effect size in that comparison – a *d* for gender – the analysis does not reveal which gender or genders are changing, only that the difference between them gets narrower with age. It may be that males are becoming more like females or vice versa, and other patterns are possible. Analyses that examine the arithmetic means for males and females can answer the question: As Figure 4 (panel c) shows, both groups increase in social dominance orientation with age, but females do so more. It should be noted that these results do not reflect data from the same samples studied across time (as was the case in the data we reported from Roberts et al.'s, 2006, analysis, panel b), but are instead a cross-sectional result.

Fixed-, random-, and mixed-effects models. The preceding material has assumed the most basic form of meta-analytic statistics, models based on fixed-effects assumptions, which are the most popular and generally match most closely the assumptions of primary-level research. *Fixed-effects* models assume that the underlying effect sizes are fixed either as a single group or else along the range of a set of moderator values. In contrast, *random-effects* models assume that each effect size is drawn at random from a universe of related but separate effects (for discussions, see Hedges & Vevea, 1998; Hunter & Schmidt, 2000; Lipsey & Wilson, 2001). In essence, to the extent that there is a lack of homogeneity in the studies' outcomes, random-effects model always provide an estimate of the population effect size whose significance is more conservative and ignores moderator dimensions and, hence, should be understood as such. Reviewers of meta-analyses commonly request random-effects models instead of fixed-effects models when the overall homogeneity statistic is significant. Yet, when the goal of the review is to assess models with moderator dimensions, fixed-effects models can be perfectly appropriate. Random-effects meta-analytic models provide only an estimate of mean effect size *without* moderators.

One way to maintain the overall random-effects assumption but also fix the effect sizes along certain moderator dimensions is to employ *mixed-effects* models, which assume that the variability in the effect size distribution is attributed to some systematic between-study differences and an additional

Size	Effect size metric		
	d	r	r^2
Small	0.20	0.100	0.010
Medium	0.50	0.243	0.059
Large	0.80	0.371	0.138

Figure 5 Cohen's (1969) Guidelines for magnitude of d and r , with r in its biserial form (see Figure 2).

unmeasured random component. It is important to note that mixed-effects models are by definition well fitting (see Overton, 1998). The risk with their use parallels the risk with purely random-effects models: they may be oversimplifications of the actual patterns present in a given literature.

Statistical power assumptions underlie all of the analyses that we have discussed (Murphy & Myers, 2004), and power will primarily vary according to the studies' sample sizes (N) and the numbers of studies included (k). If there are few studies implied in a Q value, then there is likely to be low power for assessing the assumption that the effect sizes are consistent (Hedges & Pigott, 2001). Conducting power analyses is particularly important for interpreting moderator tests and the failure may result in misleading information; if power is low, Hedges and Pigott (2004) recommended not conducting moderator analyses or including the power analysis so readers may be able to correctly interpret the outcomes of the meta-analysis.

Interpreting and reporting the meta-analytic findings

Parallel to primary-level studies, in a meta-analysis the final step concerns the interpretation of what has been found in the investigation. These results are juxtaposed against the expectations that led the review and in relation to the literature that led up to it (see the right-most column in Figure 1), as well as being careful to recognize the limitations implicit in the review. One of the most important points of consideration is how to interpret mean effect sizes. Figure 5 shows the most conventional standard. Developed by Cohen (1969, 1992), a value of $d = 0.20$ is small, 0.50 is medium, and 0.80 is large (the Figure provides comparable figures for the correlation, reflected in its biserial form). From this standpoint, Figure 3's (see panel a or b) overall gender differences in social dominance orientation are nearly medium but Figure 4 (panel a) shows that middle-aged samples exhibit a small difference. The changes in social dominance from adolescence to middle age that Roberts et al. (2006) documented (Figure 4, panel b) are large. These guidelines for effect size magnitude are just that: *guidelines*. There are some effects that, given the nature of

the phenomenon in question, can be considered larger or smaller than Cohen's guidelines would suggest, depending on expectations (Prentice & Miller, 1992) or statistical operations (McGrath & Meyer, 2006). An effect size of $d = 0.20$ would be considered quite large in some domains but not others. If the effect size concerned ratings of advertisements promoting a particular political candidate, it would probably mean that the advertisement's favored candidate would win the election. If the effect size concerned the magnitude of a mood-manipulation strategy, it would probably be considered rather trivial. Knowing how to interpret the magnitude of an effect size depends a great deal on an analyst's familiarity with the phenomenon of interest as well as on its real-world implications.

Meta-analytic reviews are structured very similarly to primary research articles and like a primary research article, should tell a good story (Cooper & Hedges, 1994). The introduction of a meta-analysis should provide a comprehensive literature review and should clearly outline the research problem of interest. If any relevant meta-analyses on the topic have previously been published, they should be referenced with an explanation of how the new meta-analysis changes or bolsters previous conclusions, if at all. The methods section of the report should clearly outline the literature search techniques, key words, and inclusion criteria implemented in the study. The results section should present the analysis strategy, statistical results, and moderator analyses, and the discussion should offer a cogent analysis of the result and orient these findings into an appropriate theoretical framework.

In short, as Figure 1 suggests, the meta-analysis write-up should address the questions most pertinent to a better understanding of the research topic. Some appropriate questions could include: now that the meta-analysis has been conducted, what do we know about this particular aspect of the social animal based on our new meta-analysis of this evidence? Have we answered the questions we raised at the outset? Have we supported particular theoretical propositions more than others? What are the strengths and weaknesses associated with this evidence? These are the types of questions that should be considered when synthesizing and presenting meta-analytic results.

The Future of Meta-Analyses Focused on the Social Animal

The information listed above identifies the most common techniques and statistics involved in meta-analysis, yet in many ways it has only scratched the surface of the process. That said, the best way to understand meta-analysis is to attempt the process oneself. There is no substitute for experience, and in terms of mastery of understanding the process of quantitative synthesis, meta-analysis is no exception. Another reason to begin to gain knowledge of meta-analysis is that new techniques and statistical tools are being developed all the time to expand the potential

of the procedure. Structural equation modeling of meta-analytic data (e.g., Albarracín, Johnson, Fishbein, & Muellerleile, 2001) is beginning to allow for the testing of causal paths after controlling for other potentially influential variables. Furthermore, meta-analyses can be an important tool to examine meta-analytic data over time, whether in terms of different samples across years (e.g., Lee et al., 2007) or for the same samples examined across time (e.g., Roberts et al., 2006). As we have illustrated, examining raw scale values as meta-analytic effect sizes also can be a valuable meta-analytic tool, and this strategy deserves much greater currency than it has received to date (see Bond, Wiitala, & Richard, 2003). Cutting-edge methods such as these make it possible to ask ever more complex empirical questions and obtain ever more nuanced models of how the social animal behaves. Furthermore, by synthesizing studies derived from a host of disciplines, meta-analyses can help to bridge different research fields in ways not previously possible, thereby facilitating interdisciplinary discourse. The meta-analyses we reviewed about dominance patterns, for example, are of import not only to social-personality psychology but also developmental and other branches of psychology, and to other disciplines, as well (e.g., public health, sociology).

Despite these advantages of meta-analytic methodology, like any scientific method, meta-analysis has limitations and its critics (e.g., Sharpe, 1997). Two common critiques are that studies included in a meta-analysis may vary in quality and that unpublished studies are underrepresented (see Bangert-Drowns, 1997; Rosenthal & DiMatteo, 2001). These complaints have primarily arisen out of a misunderstanding of meta-analytic methodology. The fact that meta-analyses have the capacity to consider all types of research is in fact one of its greatest strengths, and because factors such as quality and publication status can be included as moderators, it is possible to determine whether these factors relate to findings. Nonetheless, meta-analysis is a quintessentially correlational endeavor, with very few exceptions (Johnson & Eagly, 2000, provide further discussion). Analysts and readers alike should keep this limitation in mind.

Most importantly, no matter the subject matter, a meta-analysis must tell a good story. Otherwise, its full value may go unappreciated by readers. Although there are many paths to a good story, Sternberg's (1991) recommendations to authors of reviews are instructive: pick interesting questions, challenge conventional understandings if possible, take a unified perspective on the phenomenon, offer a clear take-home message, and write well. Thus, the practice of meta-analysis can and should embrace aspects of narrative reviewing by striving to incorporate and document the richness of the literature under consideration. Doing so not only makes meta-analysis results more accessible to other researchers but also to the public at large.

One might come to the conclusion that once a meta-analysis on a topic is published, it represents a dead end for a literature, a point beyond which

nothing more needs to be known. Interestingly, quite the opposite is true. Meta-analytic reviews have an enormous capacity for generating new empirical questions and research studies. On more than one occasion a meta-analysis has sparked renewed interest in a previously dormant research area (e.g., Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994). Furthermore, as any social scientist can tell you, scientific inquiry is based on the continual progress of research endeavors, allowing for the continual modification of the scientific knowledge base. With the passage of time, more and more studies continue to accrue, often exponentially so. Hence, as both meta-analyses and primary-level studies continue to fill journals and books (and even sneak into the mainstream knowledge base from time to time) more and more questions will develop and, as a consequence, the personality and social psychology research on the social animal will continue to expand, with meta-analysis serving as an indispensable catalyst for that growth.

Further Readings on Meta-Analysis

Cooper, H., & Hedges, L. V. (1994). *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.

Hunt, M. M. (1997). *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.

Johnson, B. T., & Eagly, A. H. (2000). Quantitative synthesis of social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 496–528). London: Cambridge University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage Publications.

Software for Conducting Meta-Analysis

Biostat Corporation (2005). *Comprehensive Meta-Analysis Version 2*. Englewood, NJ: Author.

Johnson, B. T. (1993). *DSTAT 1.10: Software for the Meta-Analytic Review of Research Literatures*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wilson, D. B. (2006). SPSS, STATA, & SAS macros for performing meta-analytic analyses. <http://mason.gmu.edu/~dwilsonb/ma.html> (last accessed August 9, 2007).

Acknowledgments

The authors thank Tania B. Huedo-Medina and Jennifer Ortiz for comments on an earlier draft of this manuscript. The preparation of this report was facilitated by US Public Health Service Grants F31-MH079759 to Marcella H. Boynton and R01-MH58563 to Blair T. Johnson.

Short Biographies

Blair T. Johnson's research is located at the intersection of social and personality psychology, health psychology, and public health; he has authored or coauthored numerous articles and chapters in these areas. He has held a fellowship from the Alexander von Humboldt Foundation and has been named a fellow of the American Psychological Association (Divisions 8 and 9) and the Association for Psychological Science. Johnson served as associate editor for *Personality and Social Psychology Bulletin* and coedited the *Handbook of Attitudes*. Dr. Johnson was coeditor of the *Handbook of Attitudes and Attitude Change* (2005, Lawrence Erlbaum Associates), has served as associate editor of *Personality and Social Psychology Bulletin*, and as a consulting editor for *Personality and Social Psychology Bulletin*, the *Journal of Personality and Social Psychology*, *Psychological Bulletin*, *Social Issues and Policy Review*, and *Vulnerable Children and Youth Studies*. Current research focuses on topics germane to social influence, including health behaviour change. Another important research focus is enhancing the repertoire of techniques that reviewers use in conducting meta-analytic reviews of scientific studies. Before coming to the University of Connecticut, where he presently teaches, Johnson taught at Syracuse University. He holds a BA in Psychology and Philosophy from Minnesota State University-Moorhead and an MS and a PhD in Social Psychology from Purdue University.

Marcella Boynton is currently a doctoral candidate in social psychology at the University of Connecticut. Her current research includes attitude-behavior relations, acculturation, HIV prevention, statistical methodology. She holds a BA in Psychology and French from Purdue University and an MA in Social Psychology from the University of Connecticut.

Endnotes

* Correspondence address: Department of Psychology, 406 Babbidge Rd Unit 1020, University of Connecticut, Storrs, CT 06269-1020, USA. Email: blair.t.johnson@uconn.edu.

¹ Note that the 'slopes' of the paired observations in this panel describe only whether one gender received greater weight than the other in the analysis. Because the variances were similar across samples, the fact that females samples tended to have more weight than male samples mainly reflects the fact that studies tended to sample more females than males.

References

- Albarracín, D., Gillette, J. C., Earl, A. N., Glasman, L. R., Durantini, M. R., & Ho, M. (2005). A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin*, **131**, 856–897.
- Albarracín, D., Johnson, B. T., Fishbein, M., & Muellerleile, P. A. (2001). Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychological Bulletin*, **127**, 142–161.

- Anderson, C. A. (2004). An update on the effects of playing violent video games. *Journal of Adolescence*, **27**, 113–122.
- Bangert-Drowns, R. L. (1997). Some limiting factors in meta-analysis. In W. J. Bukoski (Ed.), *Meta-Analysis of Drug Abuse Prevention Programs*. (National Institute on Drug Abuse Research Monograph 170, pp. 234–252). Rockville, MD: U.S. Department of Health and Human Services.
- Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, **1**, 311–320.
- Begg, C. B. (1994). Publication Bias. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 400–408). New York: Russell Sage Foundation.
- Bond, C. F. Jr., Wiitala, W. L., & Richard, F. D. (2003). Meta-analysis of raw mean differences. *Psychological Methods*, **8**, 406–418.
- Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, **119**, 111–137.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, **56**, 81–105.
- Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, **10**, 101–129.
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. *Journal of Personality*, **48**, 449–472.
- Cooper, H., & Hedges, L. V. (1994). *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.
- Cooper, H. (1998). *Synthesizing Research: A Guide for Literature Reviews* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Cooper, H., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin*, **87**, 442–449.
- Durantini, M. R., Albarracín, D., Mitchell, A. L., Earl, A. N., & Gillette, J. C. (2006). Conceptualizing the influence of social agents of behavior change: A meta-analysis of the effectiveness of HIV-prevention interventionists for different groups. *Psychological Bulletin*, **132**, 212–248.
- Duval, S., & Tweedie, R. (2000). Nonparametric 'trim and fill' method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, **95**, 89–98.
- Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods*, **6**, 161–180.
- Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods*, **10**, 444–467.
- Fiske, S. T. (2004). *Social Beings: Core Motives Approach in Social Psychology*. Hoboken, NJ: John Wiley & Sons.
- Hall, S. M., & Brannick, M. T. (2002). Comparison of two random-effects methods of meta-analysis. *Journal of Applied Psychology*, **87**, 377–389.
- Hardy, R. J., & Thompson, S. G. (1998). Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, **17**, 841–856.
- Harwell, M. (1997). An empirical study of Hedge's homogeneity test. *Psychological Methods*, **2**, 219–231.
- Hedges, L. V., & Vevea, J. L. (1996). Estimating effect size under publication bias: Small sample properties and robustness of a random effects selection model. *Journal of Educational and Behavioral Statistics*, **21**, 299–333.
- Hedges, L. V., & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.
- Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, **6**, 203–217.
- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, **9**, 426–445.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, **3**, 486–504.

- Hedges, L. V., Cooper, H., & Bushman, B. J. (1992). Testing the null hypothesis in meta-analysis: A comparison of combined probability and confidence interval procedures. *Psychological Bulletin*, **111**, 188–194.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, **21**, 1539–1558.
- Hofman, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, **31**, 1369–1385.
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods*, **11**, 193–206.
- Huedo-Medina, T., & Johnson, B. T. (2007). I^2 is subject to the same statistical power problems as Cochran's Q [Letter to editor]. *British Medical Journal*. <http://www.bmj.com/cgi/eletters/327/7414/557> (last accessed August 9, 2007).
- Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, **8**, 275–292.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.). Newbury Park, CA: Sage Publications.
- Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, **106**, 290–314.
- Johnson, B. T., & Eagly, A. H. (2000). Quantitative synthesis of social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 496–528). London: Cambridge University Press.
- Johnson, B. T., & Turco, R. (1992). The value of goodness-of-fit indices in meta-analysis: A comment on Hall and Rosenthal. *Communication Monographs*, **59**, 388–396.
- Johnson, B. T., Scott-Sheldon, L. A. J., Snyder, L. B., Noar, S., & Huedo-Medina, T. B. (forthcoming). Contemporary approaches to meta-analysis of communication research. In M. D. Slater, A. Hayes, & L. B. Snyder (Eds.), *The Sage Guide to Advanced Data Analysis Methods for Communication Research*. Thousand Oaks, CA: Sage Publications.
- Journal Citation Reports (2007). *Thompson Scientific*.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-Factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, **87**, 530–541.
- Kline, R. (2004). *Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research*. Washington, DC: American Psychological Association.
- Krippendorff, K. (1980). *Content Analysis*. Thousand Oaks, CA: Sage Publications.
- Lee, I., Pratto, F., & Johnson, B. T. (2007). *Meta-Analysis of Social Dominance Orientation Research*. Unpublished data.
- Light, R., & Pillemer, D. (1984). *Summing Up: The Science of Reviewing Research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, **48**, 1181–1209.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage Publications.
- Marcus-Newhall, A., Pedersen, W. C., Carlson, M., & Miller, N. (2000). Displaced aggression is alive and well: A meta-analytic review. *Journal of Personality and Social Psychology*, **78**, 670–689.
- McGrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: The case of r and d . *Psychological Methods*, **11**, 386–401.
- Murphy, K. R., & Myers, B. (2004). *Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods*, **3**, 354–379.
- Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, **112**, 160–164.
- Richard, F. D., Bond, C. F. Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, **7**, 331–363.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, **132**, 1–25.
- Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research* (rev. ed.). Beverly Hills, CA: Sage Publications.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, **52**, 59–82.
- Sánchez-Meca, J., & Marín-Martínez, F. (1997). Homogeneity tests in meta-analysis: A Monte-Carlo comparison of statistical power and Type I error. *Quality and Quantity*, **31**, 385–399.
- Saunders, L., Soomro, G., Buckingham, J., Jamtvedt, G., & Raina, P. (2003). Assessing the methodological quality of nonrandomized intervention studies. *Western Journal of Nursing Research*, **25**, 223–237.
- Shadish, W. R., & Sweeney, R. B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *Journal of Consulting and Clinical Psychology*, **59**, 883–893.
- Shadish, W. R., Navarro, A. M., Matt, G. E., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin*, **126**, 512–529.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, **17**, 881–901.
- Sternberg, R. J. (1991). Editorial. *Psychological Bulletin*, **109**, 3–4.
- Twenge, J. M., Zhang, L., & Im, C. (2004). It's beyond my control: A cross-temporal meta-analysis of increasing externality in locus of control, 1960–2002. *Personality and Social Psychology Review*, **8**, 308–319.
- Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, **60**, 419–435.
- Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, **115**, 323–345.