

Writing Narrative Literature Reviews

Roy F. Baumeister

Case Western Reserve University

Mark R. Leary

Wake Forest University

Narrative literature reviews serve a vital scientific function, but few resources help people learn to write them. As compared with empirical reports, literature reviews can tackle broader and more abstract questions, can engage in more post hoc theorizing without the danger of capitalizing on chance, can make a stronger case for a null-hypothesis conclusion, and can appreciate and use methodological diversity better. Also, literature reviews can draw any of 4 conclusions: The hypothesis is correct, it has not been conclusively established but is the currently best guess, it is false, or the evidence permits no conclusion. Common mistakes of authors of literature review manuscripts are described.

Narrative literature reviews form a vital part of most empirical articles, theses, and grant proposals, and of course many articles and book chapters are devoted specifically to reviewing the literature on a particular topic. Literature reviews serve a scientific field by providing a much-needed bridge between the vast and scattered assortment of articles on a topic and the reader who does not have time or resources to track them down. Reviews also present conclusions of a scope and theoretical level that individual empirical reports cannot normally address.

For individual researchers, writing a major literature review article is a very infrequent but often a very important career contribution. Yet, despite the importance of narrative literature reviews, no easy and available way to learn to write them is known. Research methods textbooks do not usually explain how to do them, even though reviewing literature is an important research method. Most graduate seminars in research methods likewise devote little or no time to them. Apprenticeship with an accomplished literature reviewer seems to be one possible strategy to learn this technique, but such specialists are rare, and moreover it is generally considered more important for stu-

dents to apprentice with someone who has expertise in empirical methods.

Most research psychologists are trained in methods of data collection, and that training generally includes how to write empirical manuscripts. Not surprisingly, researchers use their knowledge regarding empirical papers when they write literature reviews. Undoubtedly there are important similarities between writing literature reviews and writing empirical reports, and indeed Bem (1987, 1995) has emphasized such parallels and similarities in giving valuable tips about writing. Yet, important differences exist between writing empirical reports and writing literature reviews. In this article, we try to offer an overview of the special problems, advantages, opportunities, and pitfalls that pertain to narrative literature reviews, as compared with writing empirical reports.

Our own collaboration began, perhaps fittingly, with a literature review project. We had each by that point published a number of prior literature review articles and chapters. What struck us, however, as we began our work together was not how much we knew about the process, but how ignorant we still were. With each new review paper, reviewers and editors raised new and important points that improved the style and impact of our reviews. In the absence of works that explained how to write a literature review, we continued to learn on a trial-and-error, individual basis, guided by editors and their consultants. Such repetitive instruction is, of course, highly inefficient for the field, not to mention a large drain on the time and energies of the editorial board of major

Roy F. Baumeister, Department of Psychology, Case Western Reserve University; Mark R. Leary, Department of Psychology, Wake Forest University.

Correspondence concerning this article should be addressed to Roy F. Baumeister, Department of Psychology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7123. Electronic mail may be sent via Internet to rfb2@po.cwru.edu.

literature review journals, who find themselves giving the same feedback and advice over and over. We hope this article can help future writers benefit from the lessons we have learned.

We should note that this article does not deal with meta-analysis papers, which constitute an important and valuable form of literature review. Descriptions of how to do meta-analytic reviews are available elsewhere, however (Eagly, 1987; Rosenthal, 1995). We do not see meta-analysis and narrative literature reviewing as in direct competition. Where meta-analysis is usable (that is, when there are many studies available testing the same hypothesis), it is generally the preferred method. A narrative literature review is valuable, however, when one is attempting to link together many studies on different topics, either for purposes of reinterpretation or interconnection. As such, narrative literature reviewing is a valuable theory-building technique, and it may also serve hypothesis-generating functions. Meta-analysis is, in contrast, a hypothesis-testing technique. Narrative literature reviews also may be useful for testing hypotheses when meta-analysis will not work, such as when the studies are so methodologically diverse as to make meta-analytic aggregation impractical.

Goals of Literature Reviews

There are several different goals that literature reviewers may try to accomplish, and it is helpful to have one's goal clearly in mind while writing the manuscript. Five main goals can be distinguished, and these have implications for the structuring of the article and its place—or lack thereof—in the literature.

The most ambitious goal of literature review papers involves theory development. In such a paper, the author's primary objective is to propose a novel conceptualization or theory regarding some psychological phenomenon. The manuscript reviews the literature to provide a context for describing, elaborating, and evaluating the new theory, or indeed the theory may be found in the integration of the material reviewed. A slightly less ambitious but more common type of literature focuses on theory evaluation. In this type of review, the author does not offer a new theoretical perspective but rather reviews the literature relevant to the validity of an existing theory (or often two or

more competing theories). In essence, the published literature provides a database from which the author draws conclusions about the merits of existing conceptualizations.

The leading review journals most commonly publish articles that are aimed at theory construction or theory evaluation. Authors aspiring to write such reviews must therefore recognize that their task is not simply assembling and describing past work but rather is one of building or testing theory. In important respects, such an article resembles a report of a laboratory experiment: It describes empirical evidence that evaluates a theoretical hypothesis. Each piece of evidence covered in the manuscript draws its value from how it helps build or evaluate the overarching theory.

A third type of literature review surveys the state of knowledge on a particular topic. Such reviews may provide useful overviews and integrations of an area, but they are not intended to offer novel ideas, new interpretations, or sweeping conclusions. These reviews can be valuable as a means of pulling together what is known about a particular phenomenon, such as for a grant proposal, or as a resource to teachers. Because the theoretical contribution is minimal, however, the leading journals are generally reluctant to publish manuscripts of this kind.

A fourth category of literature review has problem identification as its goal. The purpose is to reveal problems, weaknesses, contradictions, or controversies in a particular area of investigation. The author may venture some tentative solutions to the problems he or she identifies but is more concerned with simply informing the field that some difficulty exists. Thus, such articles typically raise more questions than they answer, leaving it to future researchers to straighten out the mess. These would appear in journals probably more as brief articles or critiques than as full-length articles. Still, identifying problems in the empirical literature can serve a valuable scientific function.

A final, less common goal of a review article is to provide a historical account of the development of theory and research on a particular topic. Such papers are typically organized chronologically and, although their goal is primarily to trace the history of an idea, they typically provide an ongoing commentary regarding the impact and shortcomings of various contributions to the field.

Distinctive Aspects of Literature Reviews

Most research psychologists have received some training in how to write empirical reports. Not surprisingly, they use that knowledge when they write literature reviews. Undoubtedly, there are important similarities, but we focus on the differences between writing empirical reports and writing literature reviews because these offer the best opportunity to appreciate the special nature of the latter. In this section, therefore, we seek to cover what sets the literature review apart from empirical reports—in terms of both problems and opportunities.

Scope of Question and Level of Abstraction

A first point about the usefulness of literature reviews is that they allow the researcher to address much broader questions than a single empirical study can. By focusing on patterns and connections among many empirical findings, a literature review can address theoretical questions that are beyond the scope of any one study. At most, an empirical report can raise such implications in a brief and speculative way, whereas a literature review can permit conclusions about them.

Most writers of empirical articles have probably been pressured by reviewers and editors to rein in theoretical claims that were seen as excessively speculative. The editorial consultants point out, usually rightly, that the zealous author's sweeping conclusions are not warranted by his or her data set. We have two points to make about this. First, the incapacity of the data set to justify such theoretical claims probably does not reflect a flaw in that particular study, but rather it reflects a limitation in the very nature of single data sets. Any single study will rarely yield enough data that can justify broad conclusions about human nature, human behavior, or the human condition.

Second, and perhaps more importantly, those sweeping claims that the author desires to make might well be true. (After all, there are presumably some interesting and valid generalizations to be made.) A serious problem is then apparent. If such general patterns and principles exist, but no single empirical study is adequate to justify drawing such broad conclusions, how are they to be found or contributed to the field?

Another way of looking at the same problem

is to say that our empirical journals are filled with underinterpreted results. The editors are correct in insisting that a single study (or even a set of studies) does not usually permit sweeping conclusions about the human condition, especially because the scientific rules of inference require caution and parsimony. Yet the researcher in our example is sometimes correct in the belief that the empirical finding reflects an important general principle—even if the editor is correct in objecting that the finding fails to rule out other general principles or alternative possibilities at that level. The likely result is therefore that despite many empirical studies on the topic, none of them can argue the broader principle.

That, of course, is where the literature review becomes useful. Literature reviews are vital to the scientific field for bridging the gap in interpretation. Certain broad conclusions may indeed lie forever beyond the reach of any single investigation, but a literature review that examines and integrates the results of dozens of studies can address them. Without literature reviews, the field might remain permanently unable to answer some of its most fascinating questions.

Post Hoc Theorizing

Graduate students in psychology are routinely taught the importance of delineating one's hypotheses in advance (i.e., prior to collecting data). Established researchers continue to regard it as questionable and possibly unethical to theorize after one's empirical results are known. There are good reasons for such insistence, even though many experts suspect that researchers do continue to refine and develop their ideas after the data have been analyzed.

Prominent among these reasons is the danger of capitalizing on chance. Random variation will yield a fair number of spurious significant findings. If researchers were permitted to formulate their theories after completing the study, they would invent reasons for these chance findings, which would then become incorporated into the field's body of knowledge, thereby polluting it with false conclusions. Requiring researchers to formulate hypotheses in advance helps protect the field against these errors.

Such concerns do not apply to literature reviews, however. As a result, it would therefore

be unfortunate and misguided for literature reviewers to limit themselves to a priori hypothesizing in the way that empiricists should. The crucial difference is that the danger of capitalizing on chance, which can be quite serious in a single empirical study with several dozen variables and possibly hundreds of analyses, is negligible in a literature review. For a literature review to make such an error would require that dozens of studies somehow all point toward a wrong conclusion by statistical fluke, which would be highly improbable. (Other studies would almost certainly find the correct answer.) Hence it should be regarded as entirely permissible for a literature reviewer to formulate hypotheses after reading the literature. In this respect, assembling evidence by reading the literature is not the same as assembling evidence by collecting original data.

Our view carries this point a step further: It is not only permissible but positively desirable that a literature reviewer theorize after assembling the evidence. A literature review allows one to take a big step up in the level of abstraction (as compared with the level of an empirical study). When one takes that step, armed with a newly assembled set of facts and findings, one may often discover that one's original ideas formulated the issue wrongly or failed to anticipate various questions and answers. It would be foolish (and costly to the scientific field) to insist on sticking with one's original ideas. To put this another way, literature reviewers can and should remain open to new ideas far longer than empirical investigators.

Ultimately, literature reviewers have far less control over their evidence than empirical researchers because they are constrained by what other researchers have already done. To impose one's own a priori categories on that body of evidence may often be an act of procrustean rigidity that will yield misleading conclusions. Literature reviewers should allow themselves to be led by their evidence far more than empirical researchers dare. Failure to do so can cripple the capacity of a literature review to fulfill its scientific function.

The Value of Null

The pressure for positive results is another difference between writing literature reviews vs. empirical results. Authors of empirical reports

know all too well that inconclusive or null findings typically doom their chances of publication. A single study that fails to find significant results is by nature ambiguous, so editors rarely publish such papers. Knowing this, authors of empirical reports often feel pressured to make the strongest possible case that their findings contribute a clear, positive, unambiguous conclusion.

That pressure should, however, be greatly diminished for authors of literature reviews. Unlike most empirical reports, a literature review can make a useful contribution to the field by concluding that the existing data are inadequate to answer some question. Such assessments of the state of the literature help other empirical investigators know where to direct their efforts and help editors judge the novelty and importance of future empirical findings.

In extreme cases, literature reviews can make a positive contribution to the field even by concluding that no definitive answers can be drawn from the existing data, or that the null hypothesis cannot be rejected. In particular, some literature reviews tackle a set of related questions and may conclude that one or more of them remains unanswered. An empirical investigator who tested four hypotheses and learned nothing conclusive about one of them would probably be pressured to drop that one from his or her report. A literature review could conceivably be most useful for its identification of which questions remain undecided.

The value of a literature review that concludes that the evidence is inconclusive is probably most apparent when the article is in the problem identification category mentioned earlier. That is, one possible goal of a literature review is to indicate that a persistent problem or ambiguity renders some body of evidence less conclusive than is widely believed or perceived. In such cases, noting that the field knows less than it might have surmised is important. As we noted, such a goal is more compatible with briefer comments or critique articles than with full-size literature reviews aimed at developing and evaluating theory, but it nonetheless can serve a valuable corrective function.

Number of Possible Conclusions

As the previous section implies, the rules of inference regarding permissible conclusions may be quite different for empirical investiga-

tions as opposed to literature reviews. We suggest that whereas a successful empirical investigation permits only one type of conclusion, a successful literature review should permit four.

Nearly every aspiring researcher is taught that empirical studies test hypotheses that yield two possible outcomes, only one of which is meaningful. Specifically, a study either rejects its null hypothesis (while supporting its non-null hypothesis) or it fails to do so, and failure to reject the null hypothesis is deemed inherently inconclusive. The positive conclusion that accompanies the rejection of the null hypothesis is therefore the only useful conclusion an empirical study can draw. In contrast, a literature review can effectively and usefully draw any of four types of conclusion. It seems desirable that literature reviewers keep these different possible conclusions in mind and distinguish among them, instead of simply using the yes-or-no dichotomous logic that guides empirical work.

The first type of conclusion is the same as for an empirical study: The hypothesis is correct, at least based on the present evidence. A literature reviewer who finds that multiple studies provide converging evidence for the same conclusion is justified in saying that, at least given the current state of knowledge, the hypothesis is supported. Indeed, given that the literature reviewer has multiple studies and findings to work with, he or she can draw that conclusion with much more confidence than can the author of any single study.

The second possible conclusion is that the hypothesis, although not proven, is currently the best guess and should be assumed to be true until contrary evidence emerges. Such a conclusion might be reached if the evidence is subject to various flaws and biases but points consistently to the same conclusion. If all the evidence is flawed, but the flaws are different, then the most parsimonious conclusion is that the hypothesis is correct. Such a conclusion must remain tentative, however, because it is possible that a convergence of artifacts produced it.

This second type of conclusion has no analogue in empirical investigations, but it is important in literature reviews. Often one will finish reading a mass of literature fairly convinced that a hypothesis is correct but will be unwilling to assert that the findings provide definitive support for it. Logically this state of

affairs is quite distinct from asserting either that the hypothesis has been unequivocally supported or that the data are inconclusive. For researchers, this second conclusion entails that the burden of proof should at least be shifted onto other side of the argument, allowing the field to assume for the time being that the hypothesis is true (which it probably is). For practitioners and applied psychologists, such a conclusion is particularly valuable because they need the best currently available answer when a client or practical problem cannot wait many years until definitive proof may be forthcoming.

The third possible conclusion is that one does not know whether the given hypothesis is true or false. This may arise because evidence is lacking, is internally consistent and contradictory, or suffers from one or two pervasive flaws (as opposed to widely varied flaws) that render it ambiguous. As already noted, such a conclusion is usually unpublishable in an empirical report, but in a literature review it can make a valuable contribution.

The fourth possible conclusion is that the hypothesis is false. A single study with a null result is inconclusive, but if several dozen studies all fail to support a given hypothesis, then probably it is wrong. Contrary evidence can provide two distinct types of information about the wrongness of the hypothesis: The hypothesis may be wrong (as in a null-hypothesis conclusion) or the opposite of the hypothesis is correct.

This fourth type of conclusion is especially important when one is dealing with a belief or hypothesis that has become widely accepted. It is after all possible that a wrong theory will become generally accepted in any scientific field, perhaps especially a field such as psychology in which early writers put forward important conclusions with little evidence and these have been passed down without question. A single empirical study providing contrary evidence will often encounter resistance to acceptance for publication if it goes against prevailing views (and probably rightly so). A literature review of multiple findings may often be the only effective mechanism for the field to free itself of entrenched errors.

Methodological Convergence

The social and behavioral sciences address a daunting assortment of theoretical issues, and

they have developed a mind-boggling assortment of methods. In our view, this methodological diversity is a highly desirable response to a serious problem: By and large, none of the methods is perfect. The endless quest to devise new research methods reflects the perennial and valid perception that existing methods are not sure guides to the truth.

If all methods are flawed, then the conclusions are limited by the flaws. Indeed, the nature of methodological flaws is that they permit alternative interpretations of the results. The judgment about whether to publish a given empirical study often boils down to how seriously the conclusions are impaired by these methodological weaknesses and flaws.

The rule of parsimony gives the literature reviewer an important advantage over the empirical researcher, however, if the evidence being reviewed is methodologically diverse. Convergence of evidence across multiple research methods normally entails that the given hypothesis is the most parsimonious conclusion. Different flaws entail different alternative interpretations. Whereas one alternative interpretation may be enough to undermine an empirical study, it may not be able to account for all the findings if the methods are diverse.

Hence, literature reviewers are obligated to attend to the methodological diversity, and not just the quantity, of evidence. A hypothesis supported by 50 studies may in fact be shakier, more dubious, and simply more wrong than a hypothesis supported by 5 studies, if the 50 used the same method whereas the 5 used all different ones. Methodological diversity should be especially prized by narrative literature reviewers because it is something they are uniquely positioned to recognize and use in drawing conclusions about the literature.

Common Mistakes

In this section, we discuss several common mistakes that authors commit when undertaking major reviews of the literature. These are mistakes that have been pointed out to one or both of us have at various times or that we have observed when we have read literature review manuscripts written by others.

Inadequate Introduction

One common error is inadequate development of the introduction of the manuscript. Authors may skimp on presenting their conceptual and theoretical ideas early in a manuscript for many reasons. Some authors seem to think that the purpose of a literature review is simply to describe a collection of relevant findings, so no integrative theorizing is needed, and the introduction does not have any task except to convince the reader how important and interesting the topic is. (As we said, though, such reviews have little chance of being published in leading journals.) Other authors postpone the integrative theory until the discussion section after all the studies have been described, because that sequence corresponds to the author's own thought process as he or she reviewed the area—reading the evidence with an open mind and then drawing conclusions. Also, some authors may postpone their integrative “take-home message” until late in the manuscript because such a delay increases its dramatic impact. That is, they want to show what a hopeless mess the literature on their topic seemed before demonstrating how their reconceptualization can bring order to the chaos. Furthermore, as we said, empirical authors feel some obligation not to mention *post hoc* theorizing in the introduction.

For a literature review, however, it is usually necessary to present a full and vigorously integrative theoretical framework early in a manuscript. Few readers can manage to wade through 50 pages of text and dozens of facts and findings before learning what the point is. And the simple description of findings without a novel theoretical contribution is typically not sufficient to warrant publication.

We see two ways for authors to incorporate theoretical points early in a review manuscript. One is to present one's full theoretical conceptualization up front, using the remainder of the manuscript to review the literature relevant to the theory. Alternatively, an author might provide a brief “bottom-line” preview of the theory early, postponing its full elaboration until after the literature has been reviewed. Either approach provides readers with a sufficient context to make sense of the specific studies and findings cited in the review.

Inadequate Coverage of Evidence

A surprisingly common flaw in literature review manuscripts, especially those that are rejected for publication, is inadequate coverage of the cited literature. This inadequacy stems from a basic uncertainty about how much detail to give. For example, some authors cover the literature in an unbalanced fashion, devoting multiple pages to a thorough description of a few favored studies (often their own) but cursory description of the rest. As journal reviewers, we have occasionally wanted to send the manuscript back and tell the author to fill in the missing information before we could make even a tentative evaluation.

One common form of inadequacy involves citing a study's conclusion without describing the method and specific results. That is, after all, what many empirical articles do when citing literature in their introductions. Because most authors of literature reviews are empirical investigators, they may use the same style, even unwittingly, when they try their hand at writing a literature review.

A literature review loses considerable value, however, if it fails to tell the reader the nature of the evidence it presents. "X causes Y (Reference)" does not convey enough information, especially for readers who may be skeptical of the author's conclusions or who want to think for themselves. In contrast, "in a sample of A, method B produced result C (Reference), thereby supporting the view that X causes Y" is much more useful. It allows the reader to evaluate whether the conclusion fits the evidence and to understand something about the generality and methodological strength of that evidence. By neglecting to describe the nature of the evidence, the author of a literature review forces the reader to rely simply on the author's interpretation—"take my word for it." But the very purpose of a literature review is to provide a basis for accepting a conclusion without taking someone's word for it.

To put this another way: If a particular study contributes something of importance to a literature review, the review should summarize the gist of the method and results sections of that article. This does not have to be lengthy, and in fact a skilled literature reviewer can often present the relevant aspects of a study's method and results in a sentence or two. But the

evidence has to be presented at the operational level, not just at the abstract level of theoretical conclusions.

Lack of Integration

The previous section cited the problem of failing to describe the operational aspects of the reviewed studies. The opposite problem is equally destructive to the value of a literature review. This problem occurs when authors describe the procedures and observations of various studies but fail to relate them to the theoretical issues.

All reviews, regardless of their primary goal, should provide an overarching conceptualization, perspective, or point-of-view—what Sternberg (1991) called a *take-home message*—and not be content to merely recount previous ideas and research. The broader imperative is that authors of literature reviews must explain how the various studies fit together. A literature review that simply describes a series of studies on some topic has not accomplished enough to warrant publication. (Nor is it sufficient to postpone the integration until the general discussion, because most readers cannot keep all that information straight in their minds for that long.)

A literature review is primarily an integrative endeavor, and integration is best accomplished if the reader is frequently told how the individual studies fit the broad theories and patterns. To the literature reviewer, covering all the relevant studies may seem like the most important thing, but to the reader the important thing is how they fit together. To be sure, the literature reviewer should first ensure that he or she has covered the research accurately and thoroughly. But literature reviewers should also ask themselves whether they have presented each study in a way that makes its relation to the integrative themes clear and explicit.

Lack of Critical Appraisal

In the social and behavioral sciences, all conclusions are limited by the weaknesses and flaws of the evidence, and so it is essential for the literature reviewer to point out and assess those flaws and weaknesses. Often, however, authors of literature reviews neglect to do this. There are several reasons for such neglect, but

none is good. Providing a critique of the evidence is an integral, even a central part of the job of reviewing literature.

One reason authors neglect critique is that they are trying to build a case for a particular argument, and so they searched the literature and presented the results so as to make that point most convincingly. They should realize, however, that overstating the case is a serious, unnecessary, and self-defeating error in approach. It would enhance rather than diminish the value of their manuscript to indicate the weaknesses in the evidence. Moreover, helping empirical investigators see what remains to be done will probably increase the usefulness of (and future citations to) the article.

Another reason that authors of literature reviews fail to provide critiques is stylistic. Criticizing every study could possibly double the length of a manuscript, in addition to making it very tedious to read. One solution to this problem is to provide critiques of groups of studies rather than commenting on each study individually. Group or section critiques are often useful because many studies on the same topic may be subject to similar flaws and criticisms.

Ultimately, the reader of the literature review does not need to know every flaw in every study. Indeed, if one study has a flaw that is corrected in another study, the first study's flaw is rather irrelevant in the big picture. Rather, the reader needs to know how strong or how weak the overall evidence for each main point is. Group or section critiques accomplish this better than criticizing each individual study.

Another advantage of group critiques is that they permit evaluation of convergence across diverse methods, which, as we noted, is one of the advantages of literature reviews. To list a flaw in every study may give the impression that all the evidence is so weak that no conclusion can be drawn. As we noted, however, if each study has a different flaw, parsimony may dictate that the hypothesis be tentatively accepted.

Hence the most useful form of critique is normally the following: After describing the methods and results of a group of studies relevant to some point, the author should indicate briefly the major flaws in the methods and what alternative explanations they raise. Next, the consistency of the findings should be considered. Then the author should assess the

quantity and especially the methodological diversity of the evidence, keeping in mind that consistency across large quantities of methodologically diverse evidence is the best available substitute for having the proverbial direct pipeline to the truth. Finally, the author should provide a summary as to how strong the evidence is.

Failure to Adjust Conclusions

We have said that conclusions should be tempered by the flaws and weaknesses in the evidence. Sometimes authors of literature reviews will dutifully provide a critique of the evidence but then present strong, sweeping conclusions that seem to have ignored the critique entirely. As in empirical papers, authors of literature reviews sometimes go beyond the data. This pattern may be particularly common when an editor or thesis advisor insists, after reading a first draft, that critique should be added, and so the author inserts the requested critique without revising the conclusions accordingly.

Earlier, we proposed four different possible conclusions that a literature review can draw with respect to a hypothesis. It is the strength of the evidence that mostly decides among these four. The strength of the evidence encompasses its methodological rigor or lack thereof, the amount of evidence, its consistency, and its methodological diversity. In particular, these may be needed to distinguish between the two kinds of favorable judgments a literature review can make, namely whether the hypothesis is well established as correct vs. whether it is merely the best guess based on currently available evidence.

Blurring Assertion and Proof

A simpler error that is related to the failure to describe the nature of evidence is the failure to distinguish between assertion and evidence. At issue are statements of the sort, "women are smarter than men (Brown & Green, 1966)." Such statements leave unclear the crucial question of whether Brown and Green merely asserted that women are smarter or actually provided supporting evidence. In psychology, with its rich legacy of theoretical speculation based on informal observation, it is quite easy to

blur the distinction between citing someone who stated an idea and citing someone who provided evidence for it.

For writers of empirical reports, this distinction is not always important. An empirical report usually cites past literature merely to provide a context for its own evidence, which will be presented in detail in the method and results sections. In a literature review, however, the evidence is precisely in the presentation of past research, and so the nature of this evidence must be explained carefully and fully. Thus, for literature reviews, the distinction between assertion and evidence is very important.

There is undoubtedly some scholarly obligation to acknowledge important theoretical assertions by previous generations. Indeed, the bulk of the many citations to Sigmund Freud, William James, and dozens of others refer to thoughtful theoretical proposals rather than empirically demonstrated facts, at least by modern standards of empirical proof. It seems impractical to propose that citations should only be made for empirical demonstrations.

However, we recommend that literature reviewers make a persistent effort to inform the reader whether a cited source proved or merely asserted something. "Brown and Green proposed that women are smarter than men" would clearly not be confused with "in a sample of middle-aged adults, women scored higher on a test of logical reasoning than men (Brown & Green, 1966)." Because most sources cited in a literature review will presumably be empirical reports, it seems most important to be explicit in pointing out whenever some source is being cited merely for making a speculative or theoretical assertion.

Selective Review of Evidence

Although literature reviews are less subject than empirical investigations to capitalizing on chance, they are probably more susceptible to the danger of confirmation bias. Many good literature reviews involve seeing a theoretical pattern or principle in multiple spheres of behavior and evidence, and putting together such a paper undoubtedly involves an aggressive search for evidence that fits the hypothesized pattern. Areas that do not pan out are easy to ignore or skip, but the result can be a misleading impression of universality.

Selectivity in a literature review can take several forms. At worst, the author may be operating as an "intuitive lawyer" rather than in a scientist mode, in the sense of trying to make a case for one particular position or conclusion (Baumeister & Newman, 1994). This could lead him or her to cover only material that fits that view and ignore the rest. The reader is therefore left unaware of material that would weaken or contradict the argument. Such an approach is at best unfortunate and sloppy, at worst intellectually dishonest.

A less bad but still unfortunate pattern is that of selective critique. In this pattern, the author covers all the relevant evidence, both supportive and contrary to his or her view, but then applies more rigorous methodological standards to the contrary evidence than to the supportive evidence. Such thought patterns are common sources of bias in everyday thinking (Kunda, 1990; Lord, Ross, & Lepper, 1979), and literature reviewers are undoubtedly subject to them.

Ideally, a literature reviewer should spend some time searching for counterexamples or domains of evidence that would seem to contradict the main conclusions and patterns. These can be included in the article as separate sections and given the same critical appraisal as the supporting evidence. If there are important exceptions to the general patterns and conclusions, the literature review is strengthened by acknowledging them, and theory can be built further by recognizing moderators and boundary conditions. If the exceptions are merely apparent and do not on close inspection contradict the main pattern, the manuscript is strengthened by pointing them out. Thus, either way, a literature review can be improved by a deliberate search for contrary evidence.

Focusing on the Researchers Rather Than the Research

In his advice to writers of literature reviews, Bem (1995) proposed that names of researchers should always be relegated to parentheses rather than occupying grammatically prominent positions in the sentences. We find this rule excessive and can think of instances in which the sentences do need to feature the names of researchers (e.g., "Jones found one thing, but Smith found the opposite"). Nonetheless, our

disagreement with Bem is merely a matter of degree and exception, and we concur wholeheartedly with the main thrust of his argument. Good writing of literature reviews requires a concerted effort to feature the findings and ideas. Downplaying the names of researchers (such as by putting citations in parentheses) is a valuable stylistic device for ensuring that the article focuses on ideas and research rather than on theorists and researchers. It also helps the writer to avoid the appearance of making *ad hominem* arguments. As a rule of thumb, starting paragraphs in a literature review with the name of a researcher is particularly problematic, often a sign that the writer is simply describing one study after another without making a sufficient effort to integrate them.

Another exception to the guidelines of confining names to parentheses arises when a previous author is being cited for asserting a point on theoretical grounds rather than providing evidence. We have already insisted on the importance of keeping a sharp distinction between sources who said something and sources who provided evidence to support it. The most convenient way to highlight the former cases is to use precise verbs, such as "James asserted . . ." or "Johnson theorized . . ." or "Watson speculated . . ." Again, though, these are merely exceptions to Bem's rule.

Stopping at the Present

A final mistake of authors of literature reviews is to neglect to say, explicitly, what the implications for future research are. Usually this will be a subsection in the general discussion or other concluding section. Editors normally are not satisfied to publish a literature review that summarizes a large number of studies that are already published anyway, even if there are good theoretical conclusions and integrative patterns. They want a literature review to point out remaining unresolved issues and questions, if not specifically what remains to be done, as an aid and perhaps a stimulus to further research.

Such recommendations may seem minor, trivial, or obvious to the literature reviewer, but they are not. They only seem that way because the literature reviewer is so well immersed in the topic. Reviewing a body of literature puts one in the privileged position of having a broad grasp

that goes far beyond what almost anyone else (even researchers who continue to collect data on the topic) can achieve. Pointing out what remains to be studied is relatively easy from that privileged position, but quite difficult for almost anyone else. Moreover, indicating the directions for future research can help an article influence the field, both by telling researchers what they should study and by enabling future researchers to convince editors that their work is valuable and important.

Concluding Remarks

Our position is that the narrative literature review occupies a special and privileged place in the scientific enterprise. Its opportunities, epistemological constraints, and stylistic needs differ in important ways from empirical reports. By appreciating these differences, individual researchers may become more effective consumers and more successful authors of literature reviews.

References

- Baumeister, R. F., & Newman, L. S. (1994). Self-regulation of cognitive inference and decision processes. *Personality and Social Psychology Bulletin*, *20*, 3-19.
- Bem, D. J. (1987). Writing the empirical journal article. In M. P. Zanna & J. M. Darley (Eds.), *The compleat academic: A practical guide for the beginning social scientist* (pp. 171-201). New York: Random House.
- Bem, D. J. (1995). Writing a review article for *Psychological Bulletin*. *Psychological Bulletin*, *118*, 172-177.
- Eagly, A. H. (1987) *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480-498.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098-2109.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183-191.
- Sternberg, R. J. (1991). Editorial. *Psychological Bulletin*, *109*, 3-4.

Received April 11, 1997

Accepted April 17, 1997 ■