

A Guide to Writing the Dissertation Literature Review

By Justus Randolph

1. Introduction

There are many ways to derail a dissertation, and writing a faulty literature review is certainly one of them. If the literature review is flawed, the rest of the dissertation will probably be flawed. This is because "a researcher cannot perform significant research without first understanding the literature in the field" (Boote & Beile, 2005, p. 3). Experienced thesis examiners know this; in a study of the practices of Australian dissertation examiners, Mullins and Kiley (2002; as cited in Boote & Beile, 2005, p. 6) found that,

Examiners typically started reviewing a dissertation with the expectation that it would pass; but a poorly conceptualized or written literature review often indicated for them that the rest of the dissertation might have problems. On encountering an inadequate literature review, examiners would proceed to look at the methods of data collection, the analysis, and the conclusions more carefully.

Given the importance of literature reviews, it is surprising that so many of them are so poor, both in dissertations and in journal articles. Boote & Beile (2005) claim that "the dirty secret known by those who sit on dissertation committees is that most literature reviews are poorly conceptualized and written" (p. 4). But dissertations and theses are not the only types of publications that suffer from poor literature reviews. Many of the literature reviews in manuscripts submitted for publication in journals are also flawed—see Alton-Lee (1998), Grante and Graue (1999), and LeCompte, Klinger, Campbell, and Menck (2003).

Given that that so many literature reviews are poor, it is surprising that there is not more information on how to write a literature review. Boot & Beile (2005) wrote,

Doctoral students seeking advice on how to improve their literature reviews will find little published guidance worth heeding. . . . Most graduate students receive little or no formal training in how to analyze and synthesize the research literature in their field, and they are unlikely to find it elsewhere" (p. 5).

In addition to there being little information available on literature reviews, what compounds the problem is that they are very labor intensive. Gall, Borg, and Gall (1996) estimate that a decent literature review for a dissertation will take between three and six months to complete.

The purpose of this current guide is to collect and summarize the most-relevant information on how to write a dissertation literature. I begin with a discussion of the purposes of a review, I present Cooper's (1984) Taxonomy of Literature Reviews, and then discuss the steps in conducting a quantitative or qualitative literature review. This chapter ends with a discussion of common mistakes and a framework for the evaluation of literature reviews.

2. Purposes for Writing a Literature Review.

There are many practical and scientific reasons for conducting a review. One practical reason is that it is a means of demonstrating that the author is knowledgeable about the field, including its vocabulary, theories, key variables and phenomena, its methods, and its history. Another practical reason is that, with some modification, the literature review is a "legitimate and publishable

scholarly document"(LeCompte and colleagues, 2003, p. 124; as cited in Boote and Beille 2005, p. 6). Yet, another practical reason for conducting a literature is that it allows the student to find out who the influential researchers and research groups in the field are.

Besides the practical reasons for writing the review (i.e., proof of knowledge, a publishable document, and location of a research family), there are many scientific reasons for conducting a literature review. Gall, Borg, and Gall (1996) argue that the literature plays a role in:

- Delimiting the research problem.
- Seeking new lines of inquiry.
- Avoiding fruitless approaches.
- Gaining methodological insights.
- Identifying recommendations for further research.
- Seeking support for grounded theory.

In addition, Hart (1999; as cited in Boote & Beille, 2005) provides yet more reasons for reviewing the literature. These reasons include:

- distinguishing what has been done from what needs to be done;
- discovering important variables relevant to the topic;
- synthesizing and gaining a new perspective;
- identifying relationships between ideas and practices;
- establishing the context of the topic or problem;
- rationalizing the significance of the problem;
- enhancing and acquiring the subject vocabulary;
- understanding the structure of the subject;
- relating ideas and theory to applications;
- identifying the main methodologies and research techniques that have been used; and
- placing the research in a historical context to show familiarity with state-of-the-art developments. (p.13)

Another purpose for writing a literature review not mentioned above is that it provides a framework for relating new findings to previous findings in the discussion section of a dissertation. Without establishing the state of the previous research it is impossible to establish how the new research advances the previous research.

3. A Taxonomy of Literature Reviews

One way to begin planning a research review is to think about where the proposed review fits into Cooper's (1984) Taxonomy of Literature Reviews. As shown in Table 1, Cooper suggests that literature reviews can be classified according to five characteristics: *focus*, *goal*, *perspective*, *coverage*, *organization*, and *audience*. In the paragraphs that follow, I describe in more detail each of these literature review characteristics and their constituent categories.

3.1. Focus

The first characteristic deals with the focus of the review. The foci that Cooper (1988) mentions are *research outcomes*, *research methods*, *theories*, or *practices or applications*.

Literature reviews that focus on research outcomes are probably the most common type of review. In fact, the Educational Resources Information Center (1982, p. 85) defines a literature review as an "information analysis and synthesis, *focusing on findings* and not simply bibliographic citations,

Table 1. *Cooper's Taxonomy of Literature Reviews*

Characteristic	Categories
Focus	Research outcomes Research methods Theories Practices or applications
Goal	Integration (a) Generalization (b) Conflict resolution (c) Linguistic bridge-building Criticism Identification of central issues
Perspective	Neutral representation Espousal of position
Coverage	Exhaustive Exhaustive with selective citation Representative Central or pivotal
Organization	Historical Conceptual Methodological
Audience	Specialized scholars General scholars Practitioners or policymakers General public

Note. Adapted from Cooper and Hedges (1994b).

summarizing the substance of a literature and drawing conclusions from it" (italics mine). In terms of a developing a research rationale, an outcomes-oriented review could help establish that there is a lack of and a need for information on a certain research outcome, and help justify an outcome study.

While most literature reviews focus on research outcomes, other types of reviews (i.e., methodological reviews) concentrate on research methods. In a methodological review, one might investigate the research methods in the field to help inform outcomes-oriented research by identifying key variables, measures, and methods of analysis. The methodological review is also helpful for identifying the methodological strengths and weaknesses in a body of research and for examining how research practices differ across groups, times, or settings. Methodological reviews can be combined with outcome reviews to help identify how the methods used interact with the outcomes-oriented that are found. In terms of a research rationale, a methodological review might help justify the proposed dissertation research if it turns out that the previous research has been methodologically flawed.

Other types of literature reviews concentrate on theories. The review of theories can help establish what theories already exist, what are the relationships between the existing theories, and to what degree the existing theories have been substantiated. A theoretical review is necessary, for example, if the dissertation will advance a new theory. In terms of the research rationale, a theoretical review can help establish that there is a lack of theories or that the current theories are insufficient, and,

therefore, help justify that a new theory should be put forth.

Finally, some types of reviews focus on practices or applications. For example, a review might concentrate on how a certain intervention has been applied or on how a group of people tend to carry out a certain practice. In terms of a research rationale, this type of review can help establish that there is a practical need that is not currently being met.

While a dissertation review will probably have a primary focus, it will also probably be necessary to address all of the foci mentioned here. For example, a review with an outcomes-oriented focus would probably also deal with methods so that any methodological flaws that might affect an outcome could be identified and so that the methods used in the past can inform the current methods. An outcomes-oriented review would probably also deal with any theories that are related to the phenomenon being investigated and also touch upon the practical applications of the knowledge that will be gained from the dissertation.

3.2. *Goal*

The goal of many reviews is to integrate to generalize findings across units, treatments, outcomes, and settings, to resolve a debate within a field, or to bridge the language used across fields. In other types of reviews the goal is to critically analyze the previous research, to identify central issue, or to explicate a line of argument within a field.

A dissertation review will probably have multiple goals. If a dissertation will only be a review, then the author will probably be mostly interested in integration, but will also critically analyze the research, identify central issues, or explicate a argument. However, if a dissertation author is using the literature review to ground a later investigation, then the goal will probably have more to do with critically analyzing the literature—to find a weakness in it and propose to remedy that weakness with the dissertation research. The author of a dissertation review will also probably have to also identify the central issues. Also, the author will still have to integrate reviews to be able to present the reader with the big picture. Without integration, the map of the research landscape will be as big the research landscape itself.

3.3 *Perspective*

As in primary research, review authors will have to decide what perspective to take. They can take on a neutral and objective perspective and claim to just present the facts. Alternately, a review author can take a subjective perspective and discuss how the author's preexisting biases and experiences might have affected the review.

3.4 *Coverage*

Deciding on how wide of a net to cast is critical step in conducting a review. In an exhaustive review, the reviewer promises to have found every available piece of research on a certain topic, whether it was published or unpublished. Obviously, no arguments can be about incomplete coverage if an author includes every relevant piece of research that exists. However, finding every piece of research might take much more time than is available. In an exhaustive review, the key is to define the population in such a way that it is bounded and the number of articles to reviewed are manageable. This type of review Cooper (1984) calls an *exhaustive review with selective citation*. In an exhaustive review with selective citation the reviewer might choose to look at only articles published in journals, but not conference papers; however, there should be some theoretical reason to exclude conference papers. It might be the case that conference papers might paint a different, yet valid, picture of a phenomenon.

Another approach is to take a representative sample of articles and make inferences from that sample to the entire population of articles. Random sampling is often used as a method of getting a

representative sample, but random sampling is far from foolproof. Another approach is to gather evidence that shows that the representative sample is actually representative. The safest approach is to do both.

Still another approach to selecting articles is to take a purposive sample—when reviewers take a purposive sample they might only examine the central or pivotal articles in a field. Those who take a purposive sample will have to convince the reader why the articles they have chosen are the central or pivotal articles in a field and why the articles not chosen are not the central or pivotal articles in a field.

3.5 Organization

There are many ways to organize a review. Three of the most common formats are the *historical format*, the *conceptual format*, and the *methodological format*. In the historical format one organizes the review chronologically. Obviously, this is best for reviews where one wants to emphasize the progression of research methods, theories, or a change in practices over time. Another common way to organize the review is around concepts. For example, one might organize a review around the propositions in a research rationale. In a theoretical review, one could organize the review around the different theories in the literature. Finally, another way to organize the review is to use the format for an empirical paper (i.e., introduction, method, results, and discussion). In some cases one might mix and match these different methods. For example, one might begin with an introduction, method, and then present the results in a historical or conceptual format before moving on to the discussion of results.

3.6 Audience

The final characteristic in Cooper's Taxonomy of Literature Reviews is *audience*. For a dissertation, the supervisor and reviewers of the dissertation are the primary audience. The scholars within the field that the dissertation relates to are the secondary audience. Avoiding writing the dissertation literature review for a general, non-academic audience. What constitutes a good book is probably not what constitutes a good dissertation, and vice versa.

4. How to Conduct a Literature Review

Take a look at the list below. Does it look familiar? Although it could be a step-by-step guide for how to conduct primary research, in fact they are the stages for conducting a literature review (see Cooper, 1984).

1. Problem formulation.
2. Data collection.
3. Data evaluation.
4. Analysis and interpretation.
5. Public presentation.

If one thing needs to be realized about conducting and reporting a literature review it is this: **The stages for conducting and reporting a literature review parallel the process for conducting primary research.** Whatever, one knows about conducting primary research applies, with a few modifications, to conducting secondary research (i.e., conducting a literature review). One should have (a) a rationale for conducting the review; (b) research questions or hypotheses that guide the research, (c) have an explicit plan for collecting data, including how units will be chosen, (d) have an explicit plan for analyzing data, and (e) have a plan for presenting data. Instead of human participants, the units in a literature review are the articles that are reviewed. The same issues of validity and reliability that apply in primary research also apply in secondary research. And, as in primary research, the stages might be iterative and might not necessarily come in the order

presented above.

Table 2, from Cooper (1984), is a framework showing the research stages in conducting a literature review. It explains the research questions that are asked in every stage, the primary functions of each research stage, what procedural differences could lead to differing review conclusions, and the potential sources of invalidity in each stage. In the sections below (4.1 – 4.6), I discuss the steps Cooper (1984) suggests for conducting a literature review. It might also be helpful to see Table 3, a framework for evaluating the dissertation literature, to see what is expected of the end result.

4.1 Problem formation (for the literature review)

Once one has identified what type of review to conduct (using Cooper's taxonomy in Table 1), the next step is to focus the review further through formulating a problem for the review. In this problem formation step, the reviewer will decide on what questions the literature review will answer and decide in a very explicit way on what the criteria for including an article in the review are. Here I make a distinction between literature review questions (i.e., questions that can be answered by reviewing the secondary research) and empirical research questions (i.e., questions that can only be answered through primary research). The literature review is the primary source of the empirical research question.

The first step in problem formation is to create questions that guide the literature review. Those questions should be largely influenced by the goal and focus. For example, if the goal of the review were to integrate research outcomes, then a research question for the literature review might be: *From the previous literature, what is the effect of intervention X on outcomes Y and Z?* If the goal were to critically analyze the research methods in the previous literature, research questions for the literature review might be: *What research methods have been used in the past to investigate phenomenon X? What have been the methodological flaws of those methods?* If the focus of the review were on theories and the goal was to identify central issues, then a research question for the literature review might be: *What are the central theories that have been used to explain phenomenon X?* At this point, to avoid reinventing the wheel, it is necessary to search for literature reviews that have already answered these or related questions. There might have been other literature reviews that have answered these questions already.

The next step in problem formation is to explicitly determine the criteria for deciding which articles to include in the review and which articles to exclude. There are called the *criteria for inclusion and exclusion*. The criteria for inclusion and exclusion are influenced by the review's focus, goals, and coverage. Below is an example of the criteria for inclusion and exclusion in a review of the research on response cards (Randolph, 2007b):

Studies were included in the quantitative synthesis if they met each of the following criteria:

1. The study reported means and standard deviations or provided enough information to calculate means and standard deviations for each condition.
2. The use of write-on response cards, preprinted response cards, or both was the independent variable.
3. Voluntary single-student oral responding (i.e., hand raising) was used during the control condition.
4. The study reported results on at least one of the following dependent variables: participation, quiz achievement, test achievement, or intervals of behavioral disruptions.
5. The report was written in English.
6. The data from one study did not overlap data from another study.
7. The studies used repeated-measures-type methodologies.
8. For studies that used the same data as another study (e.g., a dissertation and a journal

Table 2. *The Research Stages in Conducting a Literature Review*

Stage Characteristics	Research stage				
	<i>Problem formation</i>	<i>Data collection</i>	<i>Data evaluation</i>	<i>Analysis and interpretation</i>	<i>Public presentation</i>
Research questions asked ->	What evidence should be included in the review?	What procedures should be used to find relevant evidence?	What retrieved evidence should be included in the review?	What procedures should be used to make inferences about the literature as a whole?	What information should be included in the review report?
Primary function in review ->	Constructing definitions that distinguish relevant from irrelevant studies.	Determining which sources of potentially relevant sources to examine.	Applying criteria to separate "valid" from "invalid" studies.	Synthesizing valid retrieved studies.	Applying editorial criteria to separate important from unimportant information.
Procedural differences that create variation in review conclusion ->	1. Differences in included operational definitions. 2. Differences in operational detail.	Differences in the research contained in sources of information.	1. Differences in quality criteria. 2. Differences in the influence of nonquality criteria.	Differences in the rules of inference.	Differences in guidelines for editorial judgment.
Sources of potential invalidity in review conclusions ->	1. Narrow concepts might make review conclusions less definitive and robust. 2. Superficial operational detail might obscure interacting variables.	1. Accessed studies might be qualitatively different from the target population of studies. 2. People sampled in accessible studies might be different from target population of people.	1. Nonequality factors might cause improper weighting of study formation. 2. Omissions in study reports might make conclusions unreliable.	1. Rules for distinguishing patterns from noise might be inappropriate. 2. Review-based evidence might be used to infer causality.	1. Omission of review procedures might make conclusions irreproducible. 2. Omission of review findings and study procedures might make conclusions obsolete.

Note. From Cooper (1984).

article based on the same dataset), only the study with the most comprehensive reporting was included to avoid the overrepresentation of a particular set of data. (pp. 115-116)

The criteria should be explicit and comprehensive enough so that any article that comes to light could be included or excluded based on only those criteria. Also, the criteria should be explained in enough detail that two people if given the same set of articles would more or less end up with the same subset of articles that should be included and the same subset of articles that should be excluded. In fact, in reviews where reliability is important, like when the whole dissertation or thesis is a review, researchers often test the reliability of their inclusion/exclusion system by having other individuals judge which of a given set of articles should be included and excluded and determining how close the judges are in agreement.

It has been my experience that creating a good set of inclusion/exclusion criteria takes much pilot testing through trial and error. Often there are ambiguities in the criteria and some articles "fall through the cracks." When this happens, the reviewer will have to refine the criteria for inclusion and exclusion. Recursively pilot-testing the criteria is a time-consuming process, but it is much less time-consuming than having to start over after much data have been painstakingly collected and analyzed.

4.2 Data collection

In this stage, the goal is to collect an exhaustive, semi-exhaustive, representative, or pivotal set of relevant articles. Like in primary research, the researcher of secondary data has to have a systematic plan for data collection and keep documentation of how the data were collected. The reviewer should describe the data collection procedure with such detail that other reviewers theoretically would have arrived at the same set of articles had they followed the same procedures on the same day.

Data collection often starts with an electronic search of academic bases and the Internet. (Because the relevant databases vary within fields, I will not discuss them here.) When these searches are conducted, one should record the key words and keyword combinations that were used, which databases were searched, and on what date. Also note how many records each search resulted in.

It has been estimated that electronic searches only lead to about 10% of the total articles that will comprise an exhaustive review. There are several approaches to finding the other 90%. The most effective method is to search the references of the articles that were retrieved, determine which ones seem relevant, find those, read their references and keep repeating the process until one reaches a point of saturation—a point where no new relevant articles come to light.

After an initial list of relevant articles has been compiled through electronic searches and reference searching, I suggest giving that list to colleagues and experts in the field to see if they know of any articles that should be included on the list but have not been included. I have had much success finding additional relevant articles by sending a query to the main Listserv in my field and asking the members if I have left any articles off of my list that should be there. I also strongly suggest sending the final list of potentially relevant articles to one's dissertation supervisor and reviewers to see if they know of any more articles that should be included on the list.

The data collection period can stop when the point of saturation is reached and the reviewer can convince the readers that everything that can be reasonably done to identify all of the relevant articles has been done. Of course, there will always be new articles coming to light after the data collection period has ended, but unless a new article is critically important I suggest leaving out the new articles out of the review. Otherwise, the reviewer will have to open the floodgates and start the

data collection process over.

After an initial list of potentially relevant records has been created, the reviewer will have to devise a system for separating potentially relevant from obviously irrelevant studies. For example, to determine which articles are relevant and which are irrelevant, the reviewer might read every word of every electronic record, only the abstract, only the title, or some combination. Whichever method is decided, the reviewer should keep careful documentation about the process that was undertaken. After the list of potentially relevant articles has been created, the reviewer can begin to determine which of the remaining articles meet the criteria for inclusion or exclusion. For reviews in which reliability is critical, it is common to have two or more other individuals also decide which articles meet the criteria for inclusion and exclusion to determine an estimate of interrater agreement. (Neuendorf [2002] provides a thorough discussion of methods for quantifying interrater agreement.) Once all of the relevant articles have been identified and the reviewer has determined which of those articles meet the criteria for inclusion, it is then time to begin the data evaluation stage, the subject of the next section.

4.3 Data evaluation

In this stage the reviewer begins to extract and evaluate the information in the articles that met the criteria for inclusion. This stage begins with devising a system for extracting data from the articles. The type of data that one extracts will depend on the focus and goal of the review. For example, if the focus is research outcomes and the goal is integration, obviously one will extract data about the research outcomes of each article and find some way to integrate those outcomes. In this stage, one should document the types of data that will be extracted and the process for extracting this data. Sometimes this information, because it requires so much detail, is written up separately as a coding book and coding form, and those documents are included as dissertation appendices. Other times this information is included within the main body of the dissertation.

In a coding book, a reviewer documents the process and the types of data that will be extracted from each article. If the focus of the research is outcomes, for example, one will probably have one or more variables that deal with research outcomes. However, even if the research focus is on research outcomes, one will want to extract more data than just the research outcomes. One will want to identify factors that might influence research outcomes. For example, in experimental research, these factors might include the measurement instruments that were used; the independent, dependent, and mediating/moderating variables that were investigated; the data analysis procedures used; the types of experimental controls used; and many others. The factors that are necessary to examine vary from topic to topic, so looking at previous literature reviews, meta-analyses, or coding books is critical. The coding form is an electronic document, like a spreadsheet, or a physical form on which data can be recorded for each article reviewed. A freely-downloadable example of a coding book and coding sheet that was used in a methodological review dissertation can be found from Randolph (2007a).

It is very important to carefully think through the types of data that one will extract from each article and to thoroughly pilot test the coding book. The process of extracting data from articles tends to bring to light other types of data that should also be extracted, and necessitates revising the coding book and recoding all of the articles again. Because of Murphy's Law this will happen, but pilot testing the coding book will reduce the number of times and the degree of inconvenience involved. Also, if interrater reliability is an important for a review, then one should alternately pilot test the coding book and revise it until acceptable levels of interrater reliability have been achieved.

Data about the quality of research is a variable that is commonly examined in reviews. There is a debate, however, about whether to include low quality (or invalid) articles in your review. Some, like Cooper, suggest including only high quality (valid) articles in your study. Others suggest

including both high quality and low quality studies in the review and reporting if there is a difference between the two. If there is not a difference the data can be grouped together. If there is a difference, the reviewer will probably want to report results for from the high-quality articles and low-quality articles separately.

One of the goals of most reviews is to integrate or synthesize research outcomes. To integrate research outcomes, it is necessary to find a common metric—some measure to which all of the research outcomes can be translated. In a quantitative synthesis a common metric might be, for example, the difference in proportions between control and treatment groups. If an article only presents the number of success and failures in the treatment and control groups, the reviewer will have to convert those numbers into proportions and compute the difference in those proportions—the common metric.

Like in the other stages, the reviewer has to be very specific about what data to be extracted and about the process of extraction. Whether the procedures for extracting the data are written up in a separate coding book or included within the body of the dissertation, the procedures should be written with enough detail that, actually or theoretically, a second person could arrive at more or less the same results just by following the written procedure.

4.4 Data analysis and interpretation

It is at this stage that the reviewer attempts to make sense of the data that has been extracted. If the goal is integration, the reviewer integrates the data at this point. Depending on the type of data extracted, one might do a quantitative, qualitative, or mixed-methods synthesis. See sections 5 and 6 for more information on analyzing quantitative and qualitative types of literature reviews.

4.5 Public presentation

In this stage the author needs to make a decision about what information is more important and needs to be presented and which information is less important and can be left out. In a dissertation literature review, the author can be liberal about how much information to include. Like discussed in Section 3.5, there are several methods for organizing a literature: *historically*, *conceptually*, and *methodologically*. These are the most common ways, but not the only ways.

Since it is a dissertation that is being written, the primary audience is your dissertation supervisor and the other dissertation reviewers. The secondary audience are other scholars in the field. Like I mentioned earlier, the dissertation review can be revised to meet the needs of a more general audience later.

4.6 Formulating and justifying empirical research questions

Although this stage was not included in Cooper's stages, I include it here because it is an essential part of a dissertation. The literature review, combined with the research problem, should lead to the formulation of empirical research questions. At this point, the dissertation author will need to explain, using the evidence from the review, how the dissertation makes a contribution to knowledge. The American Education Research Association (2006) lists some of the ways that new research can contribute to the existing research:

If the study is a **contribution to an established line of theory and empirical research**, it should make clear what the contributions are and how the study contributes to testing, elaborating, or enriching that theoretical perspective.

If a study is intended to **establish a new line of theory**, it should make clear what that new theory is, how it relates to existing theories and evidence, why the new theory is needed, and the intended scope of its application.

If the study is **motivated by practical concerns**, it should make clear what those concerns are, why they are important, and how this investigation can address those concerns.

If the study is **motivated by a lack of information about a problem or issue**, the problem formation should make clear what information is lacking, why it is important, and how this investigation will address the need for information. (p. 3)

5. Quantitative Literature Reviews

There are two common types of quantitative reviews: narrative reviews and meta-analytic reviews. Before the method of meta-analysis came about, almost all quantitative reviews were narrative reviews. According to Gall, Borg, and Gall (1996), narrative reviews

emphasized better-designed studies, and organized their results to form a composite picture of the state of the knowledge on the problem or topic being reviewed. The number of statistically significant results, compared with the number of nonsignificant results, may have been noted. Each study may have been described separately in a few sentences or a paragraph. (pp. 154-155)

Despite the popularity of the narrative review, they tend to be severely affected by the reviewer's subjectivity. It has been shown that the conclusions of one narrative review can differ completely from another review written by a different author, even when the articles being reviewed are exactly the same (Light & Pillemer, 1984).

Currently, meta-analytic reviews have taken the forefront. In a meta-analytic review, basically, (a) the reviewer collects a representative or comprehensive sample of articles, (b) codes those articles on a number of aspects (e.g., study quality, type of intervention used, type of measure used, study outcomes), (c) finds a common metric (e.g., a standardized mean difference effect size) that allows the study outcomes to be synthesized, and then (d) examines how the characteristics of a study covary with study outcomes.

Figure 1 is an example of a type of figure found often in meta-analysis—a forest plot—and illustrates the type of information that meta-analyses typically yield. Figure 1, from Randolph 2007b, illustrates the outcomes of 13 studies that investigated the effects of response cards on academic achievement (in this case, on quiz scores). (The triangle shows the effect and the lines show the 95% confidence intervals for the effect. The common metric is a standardized mean difference effect size called *Cohen's d*). At the bottom of the figure one can find the weighted average effect size (i.e., the integrated outcome) of all 13 studies, which is about 1.1 and means that the students scored about 1.1 standards deviations higher on their quizzes when using response cards than when not using response cards.

As one might see from Figure 1, meta-analysis is a useful way to synthesize and analyze a body of quantitative research (Cooper & Hedges, 1994a; Glass, McGaw, & Smith, 1981; Lipsey & Wilson, 2001; or Rosenthal, 1991 are excellent guidebooks for conducting meta-analyses). However, some criticisms of meta-analysis are that it is subject to publication bias (i.e., that statistically significant results tend to get published more than nonstatistically significant results) and that it is too mechanistic. Some such as Slavin (1986) wisely suggest combining meta-analytic and narrative techniques. For example, one might quantitatively synthesize each study but also give a thorough narrative description of studies that are particularly relevant.

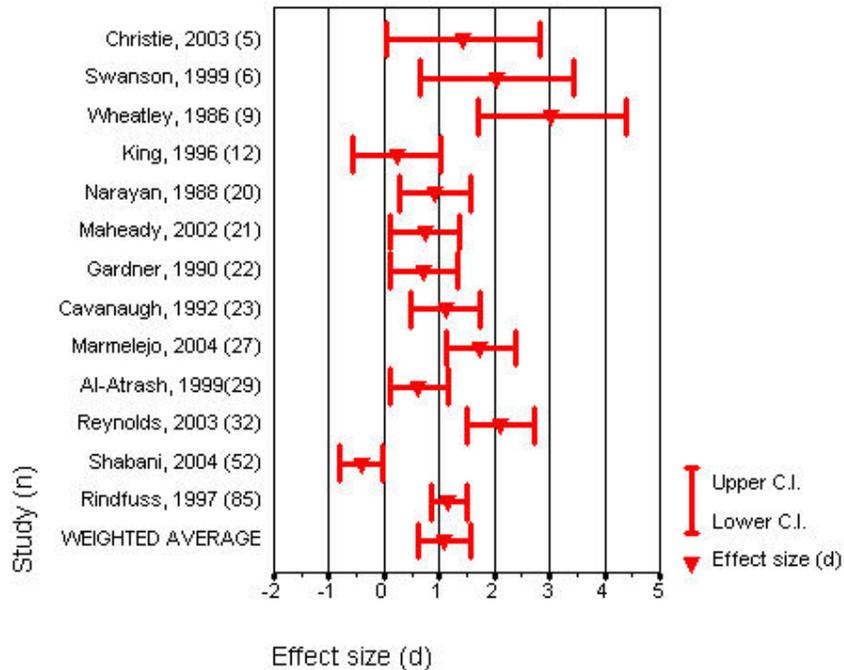


Figure 1. A forest plot of the effects of response cards on quiz achievement.

6. Qualitative Literature Reviews

Often a body of literature is primarily qualitative or contains a mixture of quantitative and qualitative results. In these cases, it might be necessary to conduct a qualitative review, either alone or as a complement to a quantitative review. In this section, I present two methods for conducting qualitative literature reviews. The first method was first put forth by Ogawa and Malen (1991). The second method, which I put forth, borrows the method of phenomenological research and applies it to conducting a literature review. Another useful resource for conducting qualitative literature reviews, but which is not described here, is Noblit and Hare (1988).

6.1 Ogawa and Malen's method

Borg, Gall, and Borg (1996) have broken down Ogawa and Malen's (1999) method into the eight steps discussed below. Note how these steps parallel the basic steps in qualitative research.

Step 1: Create an audit trail. In this step, the reviewer carefully documents all of the steps that are taken. The notion behind the audit trail is that if the review were audited, the documentation would make clear what evidence there is to support each finding, where that evidence can be found, and how that evidence was interpreted.

Step 2. Define the focus of the review. The problem formation stage mentioned in Section 4.1 is very similar to this step. In this stage one defines the constructs of the review and, thereby, one decides what to include in the review and what to leave out.

Step 3: Search for relevant literature. This step is similar to the data collection stage mentioned in Section 4.2. According to Ogawa and Malen (1999), in addition to qualitative research reports, nonresearch reports such as memos, newspaper articles, or minutes of meeting should also be included in the review and not necessarily be regarded as having less value than qualitative research

reports.

Step 4: Classify the documents. In this step the reviewer classifies the documents based on the types of data they represent. For example, some documents might be first-hand reports of qualitative research, other documents might be policy statements about the issue in question, while other types of data might be description of projects surrounding the issue.

Step 5: Create summary databases. This step is similar to the data evaluation stage mentioned in Section 4.3 above. In this stage the reviewer develops coding schemes and attempts to reduce the information in the relevant documents. On this point, Borg, Gall, and Borg (1996) wrote,

You cannot simply read all these documents, take casual notes, and then write a literature review. Instead, you will need to develop narrative summaries and coding schemes that take into account all the pertinent information in the documents. The process is iterative, meaning, for example, that you might need to develop a coding scheme, apply it to the documents, revise it based on this experience, and re-apply it (p. 159).

Step 6: Identify constructs and hypothesized causal linkages. After summary databases have been created the task is to identify the essential themes of the documents and to create hypotheses about the relationships between these themes. The goal here, unlike meta-analysis, is not to integrate outcomes and identify factors that covary with outcomes; the goal is to increase the understanding of the phenomena being investigated.

Step 7: Search for contrary findings and rival interpretations. In the tradition of primary qualitative research, it is necessary to actively search for contrary findings and rival interpretations. One might for example, reread the documents at this point to search for contrary evidence.

Step 8: Use colleagues or informants to corroborate findings. The last step in Ogawa and Malen's (1999) method, corroborating findings, also parallels primary qualitative research. In this step, one would give a draft of the report to colleagues and to informants, such as the authors of the documents included in the review, to critically analyze the review and to see if the informants agree that the review's conclusions are sound.

6.2 The phenomenological method for conducting a qualitative literature review

In phenomenological research the goal is arrive at the essence of the lived experience of a phenomenon (Moustakas, 1994). Applied as a review technique, then the goal is to arrive at the essence of researchers' empirical experience with a phenomenon. In first-hand phenomenology, the individuals who have experienced a certain phenomenon are interviewed. In using phenomenology as a review technique, the unit of analysis is the research report rather than an individual who experienced the phenomenon. Also, in using phenomenology as a review technique, the data come from an empirical research report rather than interview data.

Not surprisingly, the steps of a phenomenological review mirror the steps of phenomenological research. Those steps are briefly described below:

Step 1: Bracketing. In phenomenological research, the first step is to identify the phenomenon to be investigated. The researcher then "brackets" his or her experience with the phenomenon by telling his or her own experiences and positions with the phenomenon.

Step 2: Collecting data. The next step is to collect data about the phenomenon. In primary phenomenological research, the researcher would interview a set of people who had experience the phenomenon. In using the phenomenological method as a review tool, the reviewer would read the

reports of scientists who did research on the phenomenon. As in quantitative reviews, the reviewers still has to decide on criteria for inclusion and define the research strategy.

Step 3: Identifying meaningful statements. The next step is to identify meaningful statements. The researcher might do this by highlighting empirical claims made about the phenomenon of interest and collecting those claims, word-for-word, onto some kind of spreadsheet or qualitative software to make the data manageable.

Step 4. Giving meaning. After identifying meaningful statements, the next step is to give meanings to those statements. That is, the reviewer might put the meaningful statements into categories and then interpret and paraphrase them as groups.

Step 5. Thick, rich description. The final step is to create a thick, rich description of the essence of primary researchers' experience with the phenomenon. The goal is to describe the essence of the phenomenon as it is seen through the eyes of the researchers who investigated that phenomenon.

7. Mistakes Commonly Made in Reviewing Research Literature

In order to help avoid mistakes in conducting the literature review, I will list some of the most common mistakes here. Gall, Borg, and Gall (1996) claim that the most frequently occurring mistakes in reviewing the literature are that the researcher:

1. Does not clearly relate the findings of the literature review to the researcher's own study.
2. Does not take sufficient time to define the best descriptors and identify the best sources to use in review literature related to one's topic.
3. Relies on secondary sources rather than on primary sources in reviewing the literature.
4. Uncritically accepts another researcher's findings and interpretations as valid, rather than examining critically all aspects of the research design and analysis.
5. Does not report the search procedures that were used in the literature review.
6. Reports isolated statistical results rather than synthesizing them by chi-square or meta-analytic methods.
7. Does not consider contrary findings and alternative interpretations in synthesizing quantitative literature. (pp. 161-162)

8. Evaluating a Literature Review

Bootes and Beile (2004) have created a five-category rubric for evaluating a literature review. Those categories are *coverage*, *synthesis*, *methodology*, *significance*, and *rhetoric*. The rubric is presented in Table 3, below.

Boote and Beile (2004) used this literature review scoring rubric to rate the literature review of a random sample of 30 education-related academic dissertations. Table 4 shows a summary of their results.

How will your literature review measure up?

Table 3. *Boote and Beile's Literature Review Scoring Rubric*

Category	Criterion	1	2	3
1. Coverage	A. Justified criteria for inclusion and exclusion from review	Did not discuss the criteria for inclusion or exclusion	Discussed the literature included and excluded	Justified inclusion and exclusion of literature
2. Synthesis	B. Distinguished between what has been done in the field and what needs to be done	Did not distinguish what has and has not been done before	Discussed what has and has not been done	Critically examined the state of the field
	C. Placed the topic or problem in the broader scholarly literature.	Topic not placed in broader scholarly literature	Some discussion of broader scholarly literature	Topic clearly situated in broader scholarly literature
	D. Placed the research in the historical context of the field.	History of topic not discussed	Some mention of history of topic	Critically examined history of topic
	E. Acquired and enhanced the subject vocabulary.	Key vocabulary not discussed	Key vocabulary defined	Discussed and resolved ambiguities in definitions
	F. Articulated important variables and phenomena relevant to the topic.	Key variables and phenomena not discussed	Reviewed relationships among key variables and phenomena	Noted ambiguities in literature and proposed new relationships
	G. Synthesized and gained a new perspective on the literature.	Accepted literature at face value	Some critique of literature	Offered new perspective
3. Methodology	H. Identified the main methodologies and research techniques that have been used in the field, and their advantages and disadvantages.	Research methods not discussed	Some discussion of research methods used to produce claims	Critiqued research methods
	I. Related ideas and theories in the field to research methodologies.	Research methods not discussed	Some discussion of appropriateness of research methods to warrant claims	Critiqued appropriateness of research methods to warrant claims
4. Significance	J. Rationalized the practical significance of the research problem.	Practical significance of research not discussed	Practical significance discussed	Critique appropriateness of research methods to warrant claims
	K. Rationalized the scholarly significance of the problem.	Scholarly significance of research not discussed	Scholarly significance discussed	Critiqued scholarly significance of research
5. Rhetoric	L. Was written with a coherent, clear structure that supported the review.	Poorly conceptualized, haphazard	Some coherent structure	Well developed, coherent

Note. Adapted From Boote and Beile, 2004 (p. 9), which was adapted from *Doing a Literature Review: Releasing the Social Science Research Imagination* (p. 27), by Christopher Hart, 1999, London, Sage Publications. Copyright 1999 by Sage Publications.

Table 4. Results from using the literature scoring rubric on 30 education-related dissertation literature reviews.

Criterion	Mean (SD)
Justified criteria from inclusion and exclusion from review	1.08 (0.29)
Placed the research in the historical context of the field	2.33 (0.78)
Acquired and enhanced the subject vocabulary	2.33 (0.49)
Articulated important variables and phenomena related to the topic	2.33 (0.49)
Synthesized and gained a new perspective on the literature	1.42 (0.67)
Identified the main methodologies and research techniques that have been used in the field, and their advantages and disadvantages.	1.92 (0.79)
Rationalized the scholarly significance of the research problem	1.92 (0.79)

9. References

- Alton-Lee, A. (1998). A troubleshooter's checklist for prospective authors derived from reviewers' critical feedback. *Teaching and Teacher Education, 14*(8), 887-890.
- American Education Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*(6), 33-40.
- Boote, D. N., & Beile, P. (2004, April). *The quality of dissertation literature reviews: A missing link in research preparation*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher, 34*(6), 3-15.
- Cooper, H. M., (1984). *The integrative research review: A systematic approach*. Applied social research methods series (Vol. 2). Beverly Hills, CA: Sage.
- Cooper, H. M. (1989). Organizing knowledge synthesis: A taxonomy of literature reviews. *Knowledge in Society, 1*, 104-126.
- Cooper, H., & Hedges, L. V. (Eds.). (1994a). *The handbook of research synthesis*. New York: Sage.
- Cooper, H., & Hedges, L. V. (1994b). Research synthesis as a scientific enterprise. In H. Cooper & L.V. Hedges (Eds.), *The handbook of research synthesis* (pp. 3-14). New York: Sage.
- Educational Resources Information Center. (1982). *ERIC processing manual* (Section 5:

Cataloging). Washington, DC.

- Hart, C. (1999). *Doing a Literature Review: Releasing the Social Science Research Imagination*. London: Sage.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Education research: An introduction* (6th ed.). White Plains, NY: Longman.
- Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: SAGE Publications.
- Grant, C. A., & Graue, E. (1999) (Re)Viewing a review: A case history of the "Review of Educational Research." *Review of Educational Research*, 69(4), 384-396.
- LeCompte, M. D., Klinger, J. K., Campbell S. A., & Menke, D. W. (2003). Editor's introduction. *Review of Educational Research*, 73(2), 123-124.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied social research methods series (Vol. 49). Thousand Oaks, CA: Sage.
- Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.
- Mullins, G., & Kiley, M. (2002). "It's a PhD, not a Nobel Prize": How experienced examiners assess research theses. *Studies in Higher Education*, 27(4), 369-386.
- Neuendorf, K. A. (2002). *The content analysis handbook*. Thousand Oaks, CA: Sage.
- Noblit, G. W., & Hare, R. D., (1988) *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park, CA: Sage.
- Ogawa, R. T. & Malen, B. (1991) Towards rigor in reviews of multivocal literature: applying the exploratory case method. *Review of educational research*, 61, 265-286.
- Randolph, J. J. (2007a). *Computer science education research at the crossroads: A methodological review of computer science education research: 2000-2005*. Unpublished doctoral dissertation, Utah State University. Retrieved October 9, 2007, from http://www.archive.org/details/randolph_dissertation
- Randolph, J. J. (2007b). Meta-analysis of the effects of response cards on student achievement, participation, and intervals of off-task behavior. *Journal of Positive Behavior Interventions*, 9(2), 113-128.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Rev. ed. Newbury Park, CA: Sage.
- Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analysis and traditional reviews. *Educational Researcher*, 15(9), 5-11.