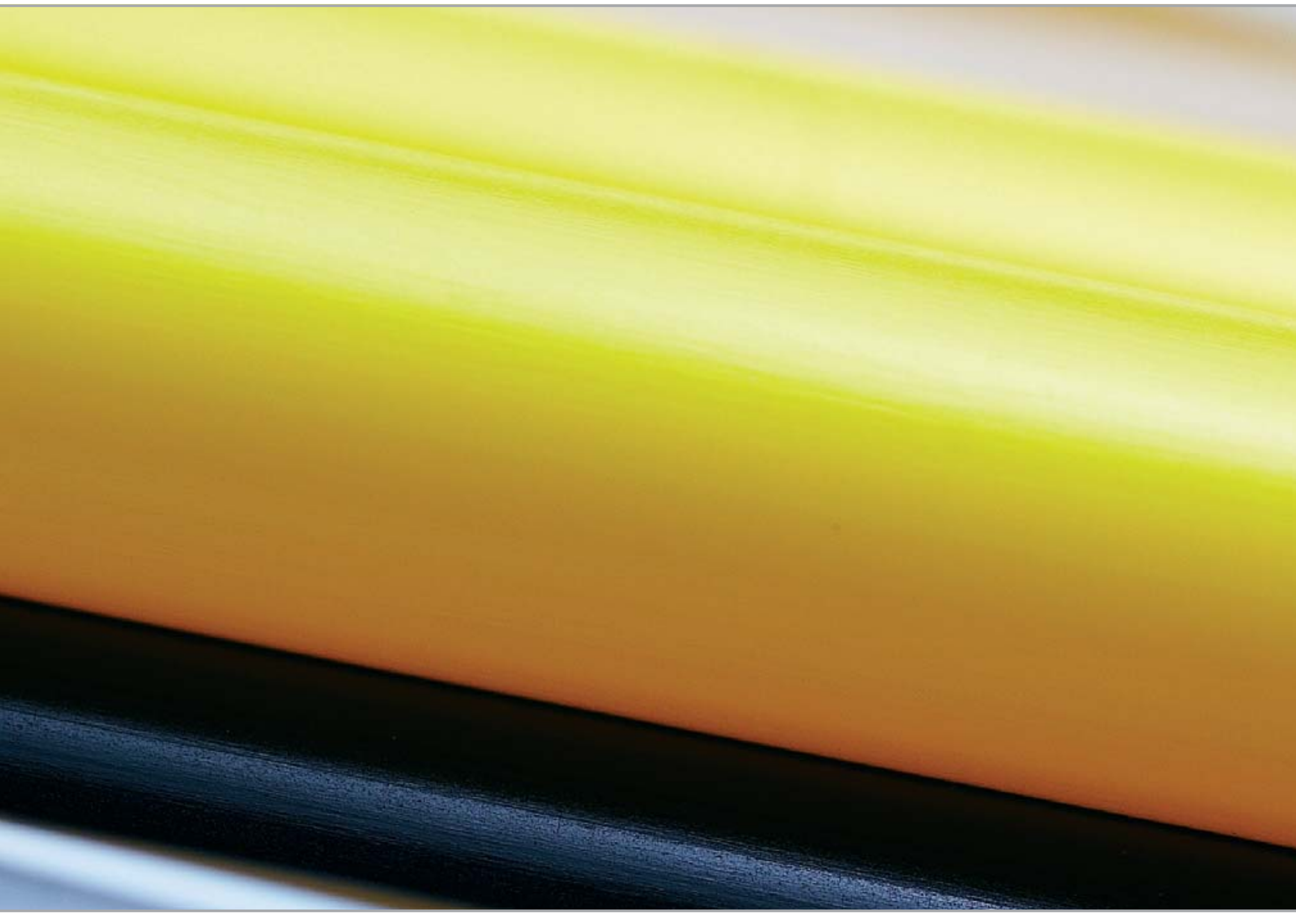
HAMK
UNIVERSITY OF APPLIED SCIENCES

Multidisciplinary Methods in Educational Technology Research and Development



Justus J. Randolph

Justus J. Randolph
Multidisciplinary Methods in Educational Technology Research and Development

ISBN 978-951-784-456-7 (PDF)
ISSN 1795-424X

HAMKin e-julkaisuja 1/2008

JULKAISIJA – PUBLISHER
HAMK University of Applied Sciences
PO Box 230
FI-13101 HÄMEENLINNA, FINLAND
tel. +358 3 6461
fax +358 3 646 4259
julkaisut@hamk.fi
www.hamk.fi/julkaisut

Book layout: HAMK Publications

Hämeenlinna, December 2008

Contents

List of Tables

List of Figures

Preface

In 2004, HAMK University of Applied Sciences began a research and development project whose goals were (a) to produce theoretical and empirical information about e-learning, (b) to develop innovative e-learning applications, and (c) to foster the growth of the e-learning enterprise in the Kanta-Häme province of Finland. To achieve those goals, a nine-element research agenda was devised. Figure 1 shows the logic of how those elements fit together to bring about the project's goals.

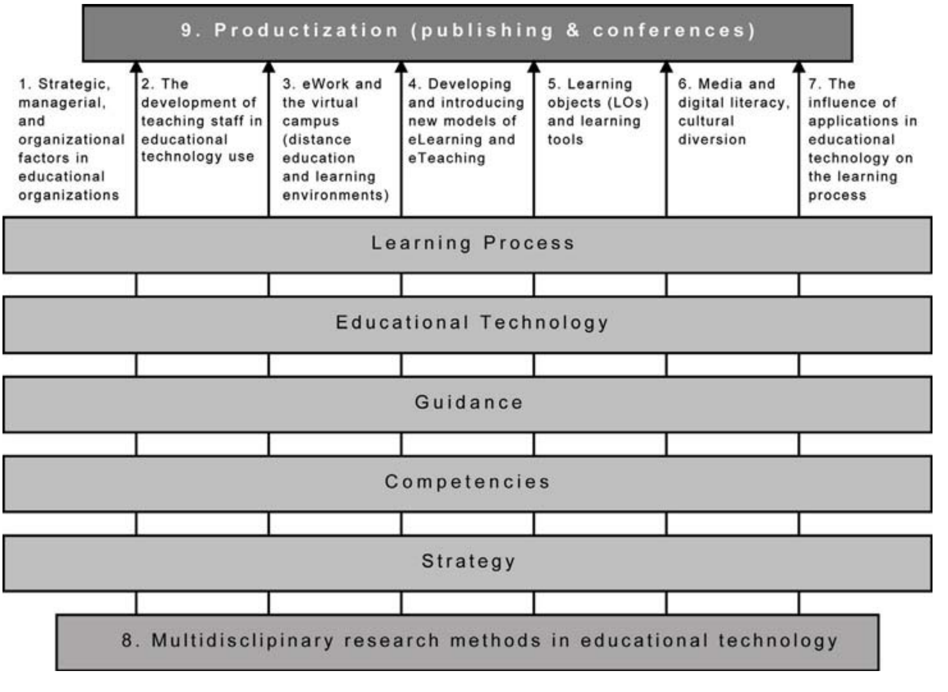


Figure 1. Strategic framework for the Digital Learning Lab research project.

The purpose of the eighth element (*Multidisciplinary research methods in educational technology*), and the purpose of this book, is to chart the multidisciplinary methods used in educational technology research and, from that charting, to produce information that can help foster improved research methods in educational technology. The improved research methods are, in turn, intended to bring about improved theoretical and empirical information about technologies for improving education.

This book is based on numerous disciplines. They include education, psychology, sociology, media studies, computing, program and policy evaluation, software engineering, business, and other disciplines whose methods can inform the field of educational technology research and development.

Organization

This book is organized into seven chapters. In the first chapter, I discuss the methods-choice debate and the factors that influence methods choice in educational technology research and development. The factors that I have identified include the following:

1. The research problem;
2. The purposes and frameworks for the research;
3. The state of and types of the previous research;
4. The type of research act implied by the research question;
5. The level of generalizability needed;
6. The level of accuracy needed;
7. The feasibility of carrying out an investigation;
8. The propriety of an investigation;
9. The utility of an investigation;
10. Whether a quantitative, qualitative, or mixed-methods tradition is being adopted;
11. Whether exploration or confirmation is needed;
12. The degree to which stakeholders participate in the research process; and
13. The degree to which the researcher becomes a participant in the intervention or setting of the investigation.

I also discuss the five major categories of research questions in educational technology research and development.

In the first part of Chapter 2, I discuss the four major quantitative research approaches (survey research, causal-comparative and longitudinal research, correlational research, and experimental research). In the second part of Chapter 2, I discuss the five major qualitative research approaches (narrative research, phenomenology, case study research, grounded theory research, and ethnography).

In Chapter 3, I provide an overview of the methodological characteristics of educational technology research by synthesizing and analyzing the results of several previous methodological reviews. The questions that the overview answers are listed below:

1. What are the meta-categories that can be used to subsume the research categories in other methodological reviews of educational technology research?
2. What are the proportions of quantitative, qualitative, and mixed methods research that educational technology researchers have tended to use?
3. How do those proportions differ over time periods and publication forums?
4. How do those proportions compare to the proportions in the field of education research proper?
5. In what proportions do educational technology researchers choose (a) research methods, (b) experimental research designs, and (c) measures?
6. How do educational technology researchers tend to report educational technology studies?
7. What suggestions are given for improving educational technology research?

In Chapter 4, I discuss the most commonly used quantitative and qualitative data collection methods in the field of educational technology research. These include the use of questionnaires, log files, tests, interviews, and direct observation.

In Chapter 5, I provide a general overview of the different methods of data analysis and refer the reader to the seminal books for each form of data analysis. I discuss the quantitative analysis of quantitative data, the quantitative analysis of qualitative data, the qualitative analysis of qualitative data, and the qualitative analysis of quantitative data.

In Chapter 6, I present information about reporting on quantitative and qualitative educational technology investigations. I provide information on writing up conventional quantitative and qualitative reports, discuss alternative styles of reporting, and present Huff's (1999) method of using exemplars.

In Chapter 7, I provide some conclusions about methods choice in educational technology research. In the Appendix, I have included a list of key questions to consider when planning educational technology research and development projects.

The Target Audiences and How They Will Benefit from this Book

The target audiences for this book are primarily educational technology students, their supervisors, and educational technology researchers. However, because of the multidisciplinarity of this text, I expect that students, instructors, and researchers from a variety of other fields could benefit from the information presented here too. In the paragraphs below, I explain how this book is expected to benefit its target groups.

The first group expected to benefit are students of educational technology research and development. With hope, the first and second chapters will help orient those student researchers to the field of educational technology research and provide them with the basic information needed to make informed decisions about what methods to choose. The third chapter will help students understand what types of methodological choices practicing educational technology researchers tend to make. The remaining chapters are intended to help students make informed decisions about data collection, analysis, and reporting and to refer them to the seminal resources for carrying out those activities. The appendix, Key Questions in Educational Technology Methods Choice, is designed to help students think through the issues that are critical when choosing methods in educational technology research and development.

The second group expected to benefit from this book are the supervisors and teachers of educational technology students. With hope, this book, in whole or in part, can serve as a course text in an educational technology research methods class and as a catalyst, focusing tool, and source of common vocabulary for academic dialogue between supervisors and their students.

The third group expected to benefit are seasoned educational technology researchers. It is my hope that the methodological factors that I have identified can serve as a starting point for clarifying the methods-choice debates in educational technology. Also, because the third chapter of this book synthesizes the research about the practices of educational technology researchers, it is my hope that it can help empirically answer questions important to the disciplinary identity of the field – questions such as:

- How do educational technology researchers and developers tend to conduct scientific inquires?
- What methods do they tend to use?
- What methods do they tend not to use?
- How do the observed practices in educational technology research differ from what is suggested as best practice, and why?

- How do the research practices in educational technology research differ from the research practices in other fields?

In addition, the answering of these questions is intended to help educational technology researchers understand the prevalent epistemological and ontological traditions in their field.

Using this Book in the Classroom

As noted earlier, one of the purposes of this book is to serve as a tool for instructing students in educational technology research methods. For example, having students read chapters one and two and then having them think through the list of key questions in educational technology methods choice at the end of this book would be a good way to familiarize students with the factors that are important in choosing appropriate methods. The key questions in methods choice at the end of this book might also serve as an intellectual organizer for supervisors and their students who are beginning to plan theses, dissertations, or projects. The third chapter can be used to introduce students to the practice of educational technology research. The fourth, fifth, and sixth chapters; which are about data collection, analysis, and reporting; can be used to familiarize students with those issues and to refer them to the essential texts in those areas. At the end of each chapter, I have included a *Questions to Consider* section that can be used as a catalyst for group or online discussions.

Positioning Myself

In the tradition of qualitative research, I think that it is worthwhile to provide the reader with some information about the author. For the past four years I have been involved with educational technology research and evaluation in various programs at the University of Joensuu and HAMK University of Applied Sciences. My main area of interest is how scientists, particularly educational technology researchers, conduct and report their research. I also am interested in research about and meta-analysis of educational interventions. My doctoral training; at Utah State University, USA; was in education research and program evaluation. I have a master's degree in education and an educational administration certification. My bachelor's degree was in English, art history, and philosophy. Although my doctoral studies were mostly quantitative in nature, I have spent the past few years developing my toolbox of qualitative methods.

Besides working for HAMK University of Applied Sciences and the University of Joensuu, I have worked as an evaluator or researcher for organizations such as the Center for Policy and Program Evaluation, The Worldwide Institute for Research and Evaluation, the National Center for Hearing Assessment and Management, Utah State University, and the Logan City School District. I also have worked as an English teacher at a private school in Poland and have had experience working as a school principal.

I was reared and educated in the United States, but I have been living in Europe for most of my adult life. I originally came to Finland, where I still reside, to do a Fulbright-sponsored research and evaluation internship.

Key Sources

Although I drew on many sources when putting together this book, I tended to draw on some more than others. Cresswell's (2007) *Qualitative Inquiry and Research Design: Choosing Among Five Traditions* was particularly useful for providing the information on the sections about the qualitative traditions and their reporting. Shadish, Cook, and Campbell's (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* was the source of much of the information on experimental designs. Also, Frechtling, Frierson, Hood, and Hughes's (2002) *The User Friendly Handbook for Project Evaluation* was drawn on heavily for the summary of data collection methods.

Some Notes on the Terms Used

Before going on, I think that it is important to clarify what I mean by some of the terms I use, because different authors use them differently. I use the term *research method* to refer to the set of steps that a researcher goes through to conduct an investigation, from designing an investigation to reporting the results of that investigation. By *research approach* I mean the approach that underlies the research method. Examples of the research approaches I discuss in this book are survey research, causal-comparative research, correlational research, experimental research, case-study research, narrative research, grounded theory research, ethnographic research, and phenomenological research. Finally by the term *research act*, I mean the types of intellectual activities that are characteristic of a certain research approach. For example, quantitative description is a characteristic intellectual activity of survey research; explanation is a characteristic activity of many types of qualitative research. I discuss the different types of research acts in Chapter 1.

Acknowledgments

This research was supported by a grant from the European Regional Development Fund admitted to HAMK University of Applied Sciences by the State Provincial Office of Southern Finland.

Thanks to Ruth Knight, Barbara Bernstein, and Byoma Tamrakar for their comments on this text. Thanks to Ilkka Jormanainen for his work on recreating several of the figures reprinted here. Also, thanks to Leena Vainio, Jorma Saarinen, Heidi Kerkola, Pasi Silander, and my other colleagues at the Digital Learning Lab for their support on this project.

A previous version of Chapter 1 was originally published as "Factors Influencing Methods Choice in Educational Technology Research," in Byoma Tamrakar's edited volume, *Modern Approaches to Digital Learning: DLL Project's Results*, published

by HAMK University of Applied Sciences, 2007. The current version of Chapter 1 was adapted with permission.

Chapter 1

Methodological Factors in Educational Technology Research and Development

The methods-choice debate is one that resurfaces with regular frequency in the education research community. This regular resurfacing is not surprising, though, given the importance of methods choice. Decisions about methods choice affect and are affected by the political, economical, and social currents of the times (Greene, Lipsey, Schwandt, Smith, & Tharp, 2007). The methods-choice debate helps determine what the research community, the media, government agencies, program funders, and the public accept as convincing evidence. And, among many other reasons, methods choice is a reflection of a research community's underlying epistemological, ontological, and axiological positions.

In the 1980s, the methods-choice debate in the social sciences flared when Cronbach et al. (1980) criticized Campbell's (1969) "reforms as experiments" view of evaluation in which laboratory research methods were favored for informing policy. Julnes and Rog (2007a, p. 1) wrote that Cronbach et al. were, "in effect nailing their ninety-five theses to the door of the edifice built on the experimental paradigm." Ten years later the debate resurfaced as what are referred to now as the *paradigm wars* – a conflict between those advocating quantitative methods and those advocating qualitative methods (see Datta, 1994; Guba, 1990; Scriven, 1993). The paradigm wars waned as the mixed-methods paradigm gained increasing acceptance. In its latest form, the methodological debate has resurfaced in response to the U.S. Department of Education's (2003) decision to give funding priority to research that adopts formal random sampling and experimental designs.

The field of educational technology research and development is not immune to these debates; in fact, the field has a long history of methodological debates of its own. The earliest artifact of this ongoing debate is Clark and Snow's (1975) initial methodological review and critique of educational technology research. Since Clark and Snow's article, at least 13 empirical methodological reviews of the educational technology literature have been conducted (see Randolph, 2007a, for a synthesis of these reviews). Other high and low points in the history of the educational technology methods-choice debate have been Phipps and Merisotis's (1999) dismissal of almost the entire body of previous research on distance learning because of its methodological flaws, Clark's (1983) criticism of media comparison studies, the treatises of the Design-Based Research Collective (2003), Reeve's (1995) criticism of the research questions of educational technology, and Williamson, Nodder, and Baker's (2001) claim that "whilst much of the literature in [the field of educational technology] is comparatively light methodologically, this can be justified by a constructivist approach to teaching and learning" (p. 1).

According to Julnes and Rog (2007a), the current methods-choice debate "is not ... about the desirability of generating evidence or about the need to consider the relative value of different methodologies. Instead, the debate is primarily over when, or under what circumstances, various methodologies provide the most useful, or actionable, evidence" (p. 2). With much wisdom, Julnes and Rog state that the way forward in the methods-choice debate is not to try to resolve the controversy, because the controversy involves deeply-rooted disagreements that are not likely to go away. Rather they suggest that way forward is "to clarify the issues to yield a more productive dialogue" (p. 2).

It is with that piece of advice in mind that I put forth the goal of this chapter: to clarify the issues, I identify and describe some of the factors that are particularly important to consider when choosing methods for educational technology research and development. To make these factors more easily understood I break them into two categories, both of which are critical to understanding methods choice: *factors that influence the formulation of the research question* and *factors that influence how a research question is answered*. The factors that influence the formulation of the research question are listed below:

- The research problem.
- The purposes of research and their corresponding traditions.
- The state of the previous research.

And, the factors that influence how a research question is answered are listed below:

- The methods used in the previous research.
- The research act implied in the research question.
- The feasibility of the research.
- Safeguards for propriety.
- The degree of utility needed.
- The degree of accuracy needed.
- The degree and kind of generalizability needed.

- The degree of stakeholder participation in the research process.
- The degree of researcher participation in the research setting.

In addition to identifying these factors, I also identify the five major types of research questions in educational technology research and development. With hope, identifying and describing these factors and research questions will help improve the productivity of the dialogue about methods choice in educational technology research within and between researchers, funders, policy makers, and practitioners.

What I do not provide here is a concrete set of rules for determining what research approach to use, what data collection methods to use, what analysis methods, or what reporting methods to use over a large set of research situations. One reason is that what may constitute the best methods choices is somewhat subjective – hence, the deep-seated disagreements about methods choice that are not likely to go away. The other reason is that while I believe that there are probably some general guidelines that apply across cases, methodological choices are heavily context-dependent. The methods that bring about actionable evidence in one setting may not bring about actionable evidence in another. Methods choice involves a careful weighing of many factors to create the most actionable evidence possible.

The methods choice factors I propose here are drawn primarily from my experience conducting methodological reviews of the educational technology literature and from key readings and previous research on methods choice (e.g., Julnes & Rog, 2007b; Shadish, Cook, & Campbell, 2002). They are also drawn from my training as a research methodologist and from my experience teaching and supervising educational technology students.

Factors Influencing the Formulation of the Research Question

Of primary importance in methods choice is the formulation of the research question because "methodology is ever the servant of substance, never the master" (Greene et al., 2007). While the research question may be of primary importance in determining the right research methods, there are a variety of factors that are of primary importance in determining the right research question – (a) the research problem, (b) the research purpose and its associated tradition, and (c) the state of the previous research. So, by substitution, the factors that are of primary importance in formulating the research question are the foundation on which methodological choices are made. (For the sake of simplicity, hereafter I use the term *research question* to refer to all of the following: scholarly research questions, evaluation questions, and development tasks.)

Before moving on to discuss the factors that influence the formulation of educational technology research problems, it might be useful to make a distinction between the different kinds of research questions. One distinction I make is between *knowledge-base questions* and *empirical research questions*. The other distinction is between *procedural research questions* and *structural research questions*.

As the name suggests, knowledge-base questions are answered through an examination of the knowledge base. The following are examples of knowledge-base questions:

- What is known about best practices in user-centered design?
- Across studies, what are the academic effects of tools that help students visualize algorithms?
- What variables are known to influence the effectiveness of educational interventions?

On the other hand, empirical research questions are meant to be answered through first-hand, empirical research. The following questions are examples of empirical research questions; however, they might also be knowledge-base questions if they had already been answered in the previous research:

- What are the effects of a new technological intervention on the long-term and short-term memory retention of vocabulary words?
- To what degree do students and teachers report that they are satisfied with a new intervention?
- In what ways do teachers and students report that a new intervention can be improved?

The utility of answering each of these types of research questions varies depending on the circumstances of the research context. Answering knowledge-base questions can, for example, help a practitioner, or a group of practitioners, find out what are the current views about best practice. Also, answering knowledge-base questions by conducting a literature review can help clarify what are the unanswered questions that are of import to the research community. On the other hand, answering empirical questions can help add to the knowledge base when there is no existing knowledge or when the existing knowledge is inadequate or in doubt. As I explain later, the choice of an empirical question is often predicated by the answer to a knowledge-base question. In addition to the distinction between knowledge-base questions and empirical research questions, a distinction can also be made between structural and procedural research questions (Stake, 1995).

To add clarity, specificity, and manageability to the general research question, it can be broken down into its structural components. For example, the general research question "What is the essence of the experience of sense of community in online courses?" could be broken down into the following structural subquestions by the issues, or structures, that are implied in the general research question:

- What do teachers experience in terms of the phenomenon of sense of community in online learning?

- What do students experience in terms of the phenomenon of sense of community in online learning?
- What medium-related contexts influence stakeholders' experience of community in online learning?

In others words, structural subquestions unpack the salient issues in a general research question.

In addition to unpacking research questions by issue, Stake (1995) suggested that research questions can also be unpacked by the research procedures to be used. For example, if one were using a phenomenological procedure to answer the general research question "What is the essence of the experience of sense of community in online courses?" the following procedural subquestions, which align with the steps in phenomenological research, might be asked:

- What statements describe teachers' and students' experiences of sense of community?
- What meanings can be inferred from these statements?
- What themes emerge from these meanings?
- What are the contexts of and thoughts about the experiences of sense of community?

In short, procedural subquestions unpack an overall research question by breaking it down into the research procedures that will be used.

By breaking down the overall research question into procedural or structural subquestions, the research process and research reporting process become clearer and more manageable. For example, the researcher can rely on the procedural subquestions to naturalistically guide the research process. Also, one often-used method of structuring the discussion section is to organize it by research subquestions. In each section of the discussion, the author answers a research subquestion by referring to the findings of the results section, relating those findings to the previous research, and explaining how those findings help resolve the research problem – the topic of the following section.

The Research Problem

My dissertation supervisor was fond of the maxim – "a lack of aspirin doesn't necessarily mean that there is a headache." Applying this notion to research, his point was that a lack of research does not necessarily mean that there is a need for research; research needs to be rationalized by both a need for and a lack of research on that topic. It is the research problem that demonstrates the need.

In this section, I make a distinction among three types of research problems in educational technology: the scholarly research problem, the evaluation problem, and the development problem. These types of research problems correspond with the different purposes of educational technology research and their associated traditions, which are discussed in the next section.

The scholarly research problem, "the intellectual quandary, dissonance, or perplexity" (Office of Research Services, 2007, p. 2) differs whether it is an applied or basic research problem, as explained below:

> *In applied research, the problem [is based on a] need, which may be based on a public policy to be fulfilled or examined and/or on data indicating some shortcoming in educational or psychological services. The need is not, however, the problem. Any one need may be the basis for a number of different research problems, depending upon the research evidence that is available and judgments about how to best address the need. For example, the need to avoid the erroneous placement of bilingual minority students in special education classes might lead to research on the sensitivity of school personnel to cultural influences on their decisions about students, on the evidence for the validity of the instruments used to classify bilingual students, or on the extent and nature of parental involvement in classification decisions. In basic research, the assumed need is for adequate knowledge, and reference to public policy or needs data is usually not necessary. (p. 2)*

The most frequently seen types of educational technology research problems (or the needs upon which they are based), which are implied by the major educational technology research questions that I discuss later, include:

- a disconnect between how educational theory informs technologies for education, and vice versa;
- a need for information about the best methods for educational technology research and development;
- a need for information about the best methods to implement and improve the utility of technological innovations;
- a need for information about the effectiveness of certain kinds of technological interventions; and
- a need for information about what factors moderate the effectiveness of certain kinds of technological interventions.

Compared to scholarly research problems, some other types of research problems, which I refer to here as *evaluation problems*, are local in scope. For example, an educational organization might have a need to respond to a local problem within their organization – perhaps there is a high degree of student attrition that needs to reduced, a need to determine if a certain distance education program should be continued or abandoned, or a need to determine if a program had been implemented as promised. Evaluation problems are typically articulated by program stakeholders.

Development problems, as the names suggests, concern the development of interventions or a lack of knowledge about how to best develop those interventions. For example, much of the field of educational technology deals with developing new or adapting existing technological interventions to solve current educational problems.

The purpose of educational technology research; whether it is scholarly, evaluative, or developmental; is to solve the types of problems mentioned above. In the next section, I discuss these different research purposes and the traditions with which they are usually associated.

The Purposes and Traditions of Educational Technology Research

Typically, research in educational technology is conducted for one or more of the following purposes:

1. to answer questions that are important for the development of an educational intervention;

2. to answer questions that are important to local stakeholders to improve, come to understand, or assign value to a program; or

3. to answer questions that are important to the scientific community.

While it is often difficult to draw clear lines between these purposes, determining the primary reason for conducting research is helpful in understanding methods choice. The research traditions that correspond primarily with the purposes of research listed above are (1) design-based research, (2) evaluation research, and (3) education research, respectively.

It is important to note that research traditions can easily overlap one another. For example, findings generated from the questions of local stakeholders might provide important insights for a scientific theory; similarly, findings from basic research might serve as a starting point for the development of an educational intervention. In the sections below, I go into more detail about each of these research traditions.

The design-based research tradition. A research tradition that has gained much credibility over the past few years and that works well for developing educational activities or tools is design-based research, which “blends empirical educational research with the theory-driven design of learning environments” (Design-Based Research Collective, 2003, p. 5). According to the Design-Based Research Collective (2003), the five characteristics of design-based research are:

> *First, the central goals of designing learning environments and developing theories or "prototheories" of learning are intertwined. Second, development and research take place through continuous cycles of design, enactment, analysis, and redesign.... Third, research on designs must lead to sharable theories that help communicate relevant implications to practitioners and other educational designers. Fourth, research must account for how designs function in authentic settings. It must not only document success or failure but also focus on interactions that refine our understanding of the learning issues involved. Fifth, the development of such accounts relies on methods that can document and connect processes of enactment to outcomes of interest. (p. 5)*

As shown above, design-based research has many characteristics, the most distinctive being its "continuous cycles of design, enactment, analysis, and redesign." In the traditional research framework, summative, generalizable, and rigorous studies are valued; however, because those types of studies are long and resource intensive, they are not feasible for the initial development of an intervention. Instead, in design-based research, numerous rapid and flexible investigations are conducted to determine how to improve an intervention. After the intervention has been perfected through many cycles of design and testing, only then does it makes sense to conduct a summative, large-scale, and resource-intensive study. What is more, design-based research is an exploratory sort of activity and, as such, can lead to insights about theories that can be later tested using confirmatory measures. In the basic form of design-based research, no particular set of methods is prescribed; the appropriate method is the one that leads to the type of information that is needed to refine the intervention.

A popular manifestation of design-based research is Bannan-Ritland's (2003) Integrative Learning Design Framework (ILD). Figure 2 shows the phases in the ILD framework and how they compare with the phases of other design traditions, such as instructional design (Dick & Carey, 1990), product design (Urlich & Eppinger, 2000), usage-centered design (Constantine & Lockwood, 1999), diffusion of innovations (Rogers, 1995), and education research (Isaac & Micheal, 1990). The ILD framework begins with an informed exploration phase that includes problem identification, a literature survey, problem definition, a needs analysis, and audience characterization. The next phase, enactment, includes researching the initial intervention design, creating a prototype, and then developing a fully detailed intervention. The next phase involves iterative cycles of pilot testing and refinement of the intervention. Bannan-Ritland describes the activities within this phase as formative testing, theory/system refinement, implementation, and evaluation. Note in Figure 2 how later stages can loop back to earlier stages in the ILD framework. For example, the results of an evaluation might indicate that the intervention needs to be redesigned. After another cycle of implementation and evaluation, it could be determined whether the refinement of the intervention had its desired effect. The final phase of ILD, evaluation of the broader impact, has to do with the dissemination, adoption, adaptation, and summative evaluation of the intervention.

Two helpful resources for design-based research are the Design-Based Research Collective's website (n.d.) and Kelly (2003). In the Design-Based Research Collective website (n.d.) one can find links to seminal writings on design-based research and links to various other design-based research resources. Kelly (2003) edited a

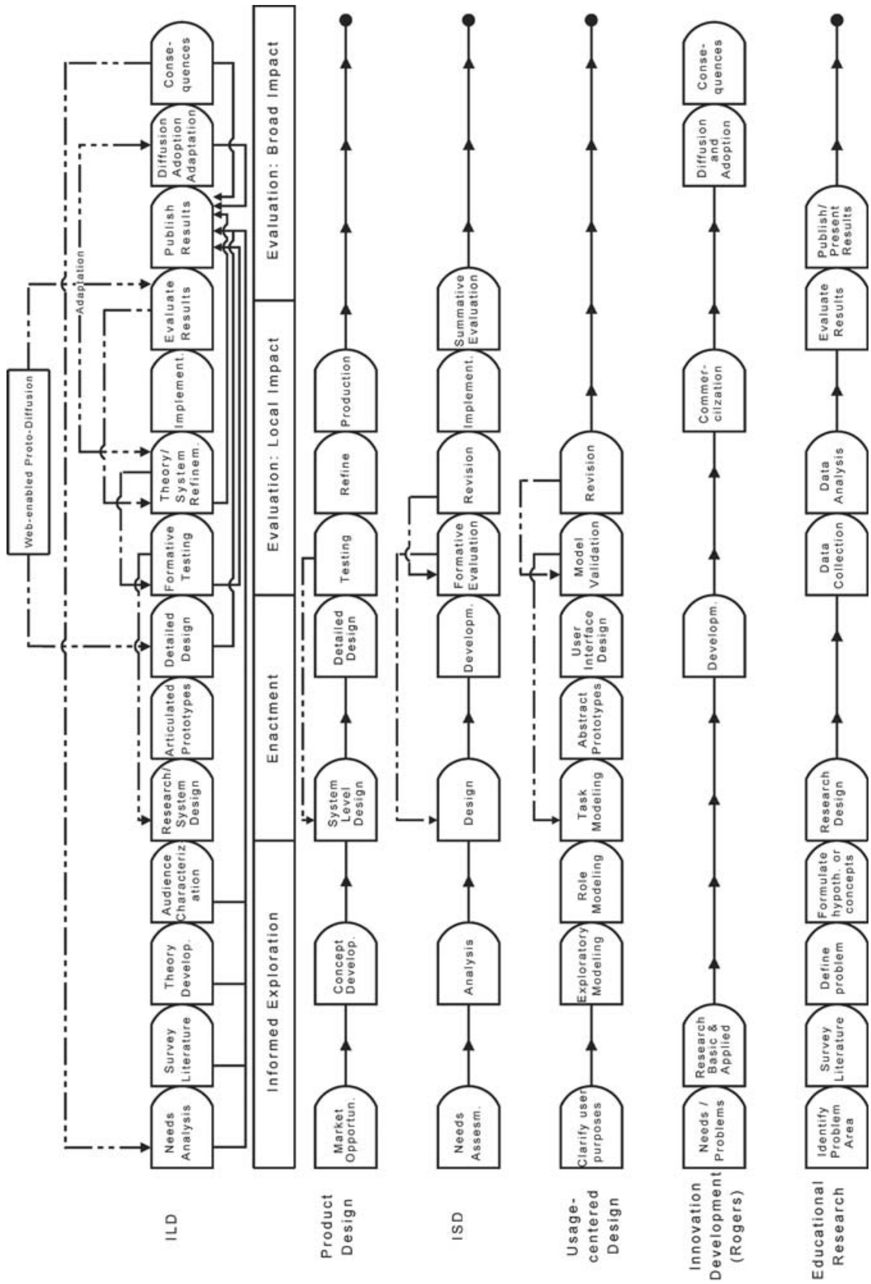


Figure 2. *Merging of design and research processes into the integrative learning design framework. From "The Role of Design in Research: The Integrative Learning Design Framework," by B. Bannan-Ritland, 2003, Educational Researcher, 32(1), p. 22. Copyright 2003 by Sage Publications, Inc. Reprinted with permission.*

special issue of *Educational Researcher* that contains a selection of articles that provide a thorough overview of the design-based research tradition.

The evaluation tradition. Three purposes are generally assigned to evaluation. Similar to design-based research, one purpose of evaluation research is to collect data that can be used to improve an intervention (formative evaluation). Another purpose is to collect data that can be used for decision-making or assigning value to a program (summative evaluation). Yet another purpose is to make sense of a program (Mark, Henry, & Julnes, 2000). Regardless of the specific purpose of evaluation, evaluation research answers questions that are primarily of interest to local stakeholders.

There are a variety of evaluation traditions to choose from, but a standard method for conducting an evaluation consists of the following steps:

- *Develop a conceptual model of the program and identify key evaluation points,*
- *Develop evaluation questions and define measurable outcomes,*
- *Develop an evaluation design,*
- *Collect data,*
- *Analyze data, and*
- *Provide information to interested audiences (Frechtling, Frierson, Hood, & Hughes; 2002, p. 15).*

There are a many good resources for evaluation research. For example, the U.S. National Science Foundation has created a series of useful, free, and practitioner-oriented evaluation handbooks. The latest in the series is Frechtling, Frierson, Hood, and Hughes's (2002) *The User-Friendly Handbook for Program Evaluation*. It provides an overview of the types of evaluation, the steps involved in conducting an evaluation, an overview of quantitative and qualitative methods, and a section on strategies for culturally responsive evaluation. Other handbooks in this series include *The User-Friendly Handbook for Program Evaluation: Science, Mathematics, and Technology Education* (Frechtling, Stevens, Lawrenz, & Sharp, 1993) and *The User-Friendly Handbook for Mixed Methods Evaluation* (Frechtling & Sharp, 1997). Seminal books in evaluation research include Herman (1987); Mark, Henry, and Julnes (2000); Patton (1990); Preskill and Torres (1999); and Weiss (1998).

The education research tradition. The final tradition I deal with here is the education research tradition. While design-based research and evaluation research may indeed be types of research on education, I have chosen to use the term *education research* to refer to research that answers questions that are of interest to the education research community. Although design-based research and evaluation research can do much to answer the questions of the scientific community, that is not their primary function.

There is no shortage of high quality books and resources on the practice of education research. They are too numerous to describe here, but I do recommend Gall, Borg, and Gall (1996) as an introductory guide to the multifaceted literature on education research.

The State of the Previous Research

For many reasons, becoming familiar with the state of the previous knowledge on a research topic, by doing a literature review, is a critical factor in one's formulation of a research question. First, conducting a literature review or needs analysis makes it possible to determine how answering one's research question will contribute to pre-existing knowledge. The American Education Research Association (2006) suggests that research can contribute to knowledge in the following ways:

- *It can contribute to an already established theory or line of empirical research,*
- *It can help establish a new theory,*
- *It can meet a practical need, or*
- *It can make up for a lack of needed information about a problem or issue. (p. 34)*

For example, the literature review should make it possible to determine whether there are established theories already and to what degree they have been substantiated. Or, from an empirical research point of view, a literature review can show if the key elements or variables have been identified, whether the associations between those elements are understood, and whether the causal mechanisms underlying the phenomenon have been identified. At any rate, these aspects about the state of the previous research will have considerable impacts on the focus of the current research. In some sense, the literature review is the mother of the research question.

Second, the literature review provides a basis for comparing and contrasting current findings with previous findings. Comparing and contrasting current findings with previous finding helps build an evidence base, puts the current study in context, and helps establish the degree to which a finding holds true over different participants, treatments, outcomes, and settings. Comparing and contrasting current and previous findings also give the current findings more meaning.

Finally, finding out about the previous research on a topic can help researchers locate themselves in what I call a *research family* and get a clear picture of how their research fits into a *research lineage*. By *research family,* I mean the individual researchers or groups of researchers that investigate the same topic. By understanding how one fits into a research family, it is easier to understand what Becher (1989) calls the *tribes and territories* of one's field. By *research lineage*, I mean the historical line of research on a particular topic. By understanding the history of research on a topic, researchers can appreciate how their research fits into that history, identify what is needed at the present, and predict what will be needed in the future. For this, Creswell (1994) suggests making a *research map* – a visual representation of how one's research fits in with the previous literature – to understand one's research lineage. An example of an outstanding research map, which was created by a PhD candidate in the field of computer science and digital information, is presented in Figure 3. It shows the relationships among the previous research studies and the relationships between the previous research and the various needs that the PhD candidate's proposed research is meant to address.

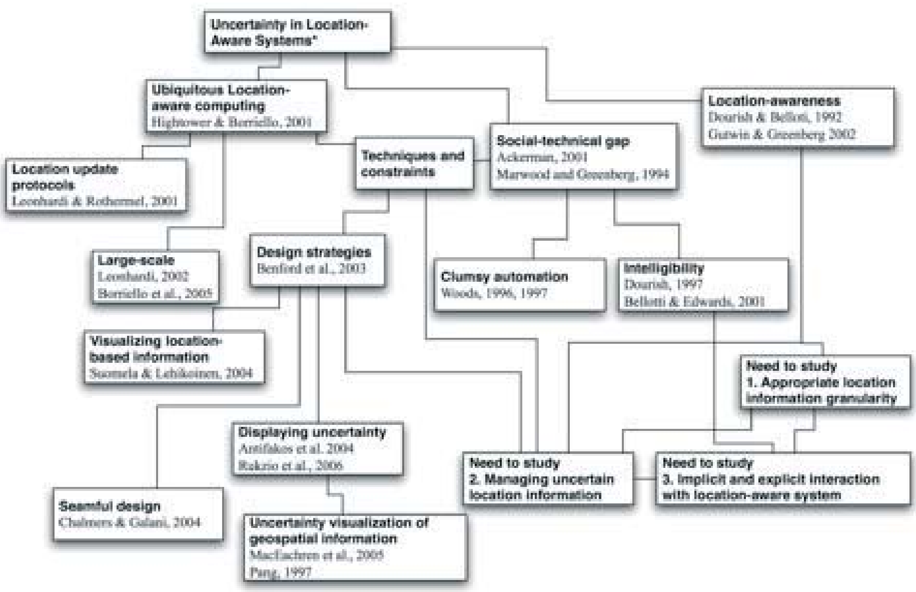


Figure 3. *Example of a research map. By F. Girardin, 2007. Copyright 2007 by F. Girardin. Reprinted with the permission of the author.*

The Five Major Categories of Educational Technology Research Questions

In the sections above, I discussed, in general, the factors that go into choosing research questions. In the section below, I discuss what types of research questions those factors tend to yield in the field of educational technology. Here I used an empirical approach to identify the major categories of research questions in educational technology, between and within the design-based, evaluation, and education research traditions. With hope, identifying and describing the categories of questions that are often seen in educational technology research will help add clarity to the debate about which methods are appropriate for answering these kinds of research questions.

Design-based research questions. The Design-Based Research Collective (2003, p. 5) has given some suggestions for the categories of research and development questions that are of critical importance. The list below summarizes the major research and development questions mentioned there.

1. Research questions that deal with the development of theories or "prototheories" of learning.
2. Research questions that deal with the interactions of an intervention and the authentic setting.
3. Research questions that deal with how an intervention causes the desired outcomes.
4. Development questions that deal with how an intervention can be improved.

In terms of Bannan-Ritland's Integrative Learning Design Framework (2003), these categories of research questions mentioned above take on specific meanings through the steps in the informed exploration stage. Those steps are *problem identification, literature survey, problem definition, needs analysis*, and *audience characterization*.

Evaluation questions. Remember that one of the primary purposes of evaluation research is to answer questions that are important to program stakeholders. So, it is no surprise that questions in evaluation research come from people who are involved in a program or intervention. Typically, evaluation questions are generated in two phases – a divergent question phase and a convergent question phase. In the divergent question phase the evaluator collects an unedited list of research questions from the people involved in the program – for example, from the administrators, practitioners, and clients. In the convergent phase, the evaluator and sometimes the stakeholders decide which of the questions from the divergent list need to be answered first and which can be answered later.

Because there has been no review of the questions in educational technology evaluation reports, a lateral review of the questions in computer science education evaluations might provide some insight into the categories of questions that educational technology evaluators strive to answer. I make the assumption here that the body of computer science education research is more or less generalizable to the body of educational technology research for two reasons: first, because so much emphasis is put on educational technology research and development in the field computer science education and, second, because the two fields exhibit many similarities in the quantity and quality of research methods used (see Randolph, 2007a).

In Randolph (2007b), I conducted a review of 29 Kindergarten through 12th grade evaluation reports of computer science education programs that had been published before March 2005. I inferred the evaluation questions from those evaluation reports. For example, if an evaluator had examined student achievement as an outcome, then I assumed that at least one of the evaluation questions had to do with the ability of the program to bring about student achievement. The factors that were examined are equally telling. For example, if gender had been examined as a factor, then I assumed that there was an evaluation question about whether the program had a differential effect for male or female participants. At any rate, Table 1 shows that the outcomes that the evaluation questions most often dealt with, in decreasing order of frequency, were *attitudes, enrollment*, and *achievement in core courses* and that the interaction factors that were examined most often were *gender, aptitude*, and *race/ethnic origin*.

Table 1. The Question Topics in K-12 Computing Education Program Evaluations.

Question #	*Question topic*	*Frequency (%)*
	Outcome (out of 67 outcomes in 19 cases)	
1	Stakeholder attitudes	17 (25.4)
2	Enrollment	13 (19.4)
3	Achievement in core subjects	14 (20.9)
4	Computer science achievement	9 (13.4)
6	Teaching practices	5 (7.5)
7	Intentions for future CS jobs/courses	3 (4.5)
8	Program implementation	2 (3.0)
9	Costs and benefits	2 (3.0)
10	Socialization	1 (1.5)
11	Computer use	1 (1.5)
	*Factors (from 19 cases)**	
12	Gender	3 (15.8)
13	Aptitude	3 (15.8)
14	Race/ethnic origin	5 (26.3)

* More than one factor was possible per case. From "*A Methodological Review of Program Evaluations in K-12 Computer Science Education*," by J. J. Randolph, 2007, manuscript submitted for publication.

Education research questions. In this section, I present the categories of research questions that have been of import to the educational technology community over the last ten or fifteen years. With hope, examining these questions of the past can help give more meaning to the research questions of the present.

In Tables 2 through Table 4, I summarize the results of three empirical reviews of the questions asked in educational technology research articles (Burghar & Turns, 1999a; Burghar & Turns, 1999b; and Burghar & Turns, 2000). In those reviews, Burghar and Turns used an emergent coding technique to create an initial set of research question categories from all the articles published over a two to four year time period from three major educational technology forums – the proceedings of *Frontiers in Education* (FIE), *Educational Technology Research & Development* (ETR&D), and *Human-Computer Interaction* (HCI). Articles published between 1997 and 1999 were selected from FIE and ETR&D; articles published between 1995 and 1999 were selected from HCI

Table 2. Major Categories of Research Questions from FIE (1997–1999).

Question #	*Question category*
1	What techniques can be used when designing technology-oriented distance learning applications?
2	What techniques can be used when designing educational technology applications?
3	How can educational technology be implemented?
4	How do students interact with educational technology?
5	How can educational technology support collaboration?
6	How can we assess student learning?
7	How can we assess the effectiveness of educational technology?
8	How have instructors at other locations used technology in the teaching of a particular subject?
9	What applications have other instructors designed to teach their subjects?

Note. The information here is summarized from J. Burghar and J. Turns, 2000.

Table 3. Major Categories of Research Questions from ETR&D (1997–1999).

Question #	*Question category*
1	How can we theoretically understand educational technology?
2	How can theory be applied to educational technology?
3	What are the effects of a given technology on practice?
4	What factors affect the implementation of a technology?
5	How can the development process be improved?

Note. The information here is summarized from J. Burghar and J. Turns, 1999a.

Table 4. Major Categories of Research Questions from HCI (1995–1999).

Question #	*Question category*
1	What methods can researchers use as they explore a design context?
2	How can user tasks be modeled and analyzed?
3	How can developers integrate users into their designs?
4	How do we characterize and design for group processes?
5	How do users interact with hypertext?
6	What can we learn about a task by studying users with varying levels of experience?
7	How can interface modalities be tailored to meet user needs?
8	How can the development process be improved?
9	How can user cognitive activity be represented in models and theory?

Note. The information here is summarized from J. Burghar and J. Turns, 1999b.

Comparing and Contrasting Questions across Forums and Traditions

Several differences across the research questions between forums and research traditions exist. First, evaluation questions tend to center more on program effectiveness and its moderators than the research questions in design-based research or in forums like FIE, HCI, or ETR&D, whose questions deal more with methodological and theoretical issues within their fields. Second, the questions in HCI seem to be more specific than the questions in design-based research or in educational technology forums like FIE or ETR&D. For example, in ETR&D a major question is "What factors affect the implementation of a technology" whereas in HCI that question is usually broken down into its sub questions – for example, "How can interface modalities be tailored to meet user needs?" Third, it appears that the questions in ETR&D are more theoretical in nature than the questions in other forums. Two out of five question types in ETR&D deal with theory – "How can we theoretically understand educational technology?" and "How can theory be applied to educational technology?"

While there are some differences in research questions across the traditions and forums, nonetheless, there is enough similarity that overall categories of research questions across traditions and forums clearly emerge. By synthesizing the questions across the different reviews of research questions presented earlier (i.e., Design-Based Research Collective, 2003; Randolph, 2007b; Burghar & Turns, 1999a, 1999b, and 2000), it appears that the questions in educational technology can be grouped into five major categories. The evidence table (Table 5) below shows the major types of educational technology research questions and the sources of the sub questions on which they were based.

Table 5. The Five Major Categories of Educational Technology Research Questions.

Category	*Source*
Questions about theories and practice	Burghar & Turns (ETR&D), 1999a, questions 1 & 2 Burghar & Turns (HCI), 1999b, question 9
Questions about research & development methods	Burghar & Turns, (ETR&D), 1999a, question 5 Burghar & Turns (HCI), 1999b, questions 1, 2, 3, & 8 Burghar & Turns (FIE), 2000, questions 1, 2, 6, & 7
Questions about technology implementation	Burghar & Turns (FIE), 2000, question 3 Randolph, 2007b, question 7
Questions about the effectiveness of an intervention	Burghar & Turns, (ETR&D), 1999a, question 3 Burghar & Turns (HCI), 1999b, question 5 Burghar & Turns (FIE), 2000, questions 4 DBRC, 2003, questions 3 & 4 Randolph, 2007b, questions 1 through 10
Questions about the factors that moderate the effectiveness of an intervention	Burghar & Turns, (ETR&D), 1999a, question 4 Burghar & Turns (HCI), 1999b, questions 4, 6, & 7 Burghar & Turns (FIE), 2000, questions 5, 8, and 9 DBRC, 2003, question 1 Randolph, 2007b, questions 11, 12 and 13

Note. DBRC = Design-Based Research Collective. The question number refers to the question # columns in Table 2 through Table 4.

Questions about theory and practice. These types of questions deal primarily with how educational and psychological theories can inform educational technology practice and how educational technology practice can inform those theories. These types of questions also include theoretical questions about the disciplinary identity of educational technology. Two hypothetical questions in this category are given below:

- How has the theory of active student response been implemented in educational technology interventions?
- Do educational technology interventions that include active student response lead to increased academic performance, as the theory suggests?

Questions about research and development methods. These types of questions deal primarily with the conduct of educational technology research and development. They deal with the methods that can be used for conducting educational technology research and development and how those methods could be improved. Three hypothetical examples of research questions in this category are provided below:

- What research methods do educational technology researchers tend to use?
- Under what circumstances do they use those methods?
- What are the strengths and weaknesses of using those methods under a variety of different research situations?

Questions about the implementation of technology. Two of the reviews presented here involved sub questions that deal with the implementation of technology. Some hypothetical examples in this question category are given below:

- What factors help increase the likelihood that a teacher will adopt an educational intervention?
- What factors help increase the likelihood that a student will adopt an educational intervention?

Questions about the effectiveness of a technological intervention. This group of questions includes formative questions about how to improve an existing technology and summative questions about how well an existing technology works in effecting a given outcome. Some hypothetical examples in this question category are given below:

- Does our educational technology intervention cause increased academic achievement?
- Does educational intervention X or Y lead to greater academic achievement?

Questions about factors that moderate the effectiveness of a technological intervention. While the previous group of questions deals with the main effects of a technological interaction, this group of questions deals with the factors that moderate the effectiveness of an intervention. Some of the factors that are examined in these questions deal with group versus individual learning, the academic subjects involved, the type of technological intervention used, the setting of the instruction, the level of previous experience, gender, age, and so on. Some hypothetical examples in this question category are given below:

- Do students who have more previous experience with computers gain more from using the educational technology intervention?
- Do the previous results concerning an intervention generalize when the intervention is used in a different setting?

Some Caveats

These categories of research questions come from articles that were written between 1995 and 2005; therefore, they reflect the state of research between 2 and 12 years ago. Naturally, the field will have progressed and some of these categories of research questions will have changed. Some of the questions will have been answered and new questions will have replaced them. These categories of research questions are only meant as a guide for situating and evaluating a set of current research questions by examining the research questions and traditions of the past.

Factors That Influence How a Research Question is Answered

Earlier I identified some of the factors that were critical in formulating a research question. Those factors included (a) the research problem, (b) the purposes and associated traditions of the research, and (c) the state of the previous research. I also identified the general types of question topics being asked in educational technology research. In this section, I discuss the factors that are important to consider when choosing methods to answer a research question once it has been formulated. Those factors include (a) the methods used in the previous research, (b) the research act implied in the questions, (c) and some salient dimensions in methods choice, such as the level of accuracy, utility, propriety, and feasibility of an investigation.

It is important to note that the factors that influence the formulation of a research question interact reflexively with the factors that influence the methods used to answer that question. For example, one might have to modify a research question if it is not feasible or if it can only be answered through an investigation that causes excessive harm to participants. While it is true that the nature of the research question implies what type of research methods are appropriate, the factors that influence how research can be carried out can limit the type and scope of research questions that can be answered.

The Methods Used in the Previous Research

The research methods and procedures used in previous research can be an invaluable guide to designing research. The previous research will show which methods have worked well in the past and which have not worked so well, which variables are important to examine and which can be left out, and what contextual and environmental factors need to be taken into account. What is more, if it is important to accumulate evidence in a field, then a researcher might want to use the methods that were used in the past so that it is easier to make comparisons across studies. Finally, a researcher may decide to make a contribution to the field by investigating a topic using a method that has not been used before. Anyway, one has to be knowledgeable about a tradition in order to break with it. After all, “the accumulated past is life’s best resource for innovation” (Brand, 1999, p. 15).

Research Acts Implied in the Research Question

In order to be able to link research questions to research methods it might be helpful to review the categories of research acts (i.e., the types of actions one takes while doing research) that are implied by the research question. Some authors call these *the purposes of research*, but I call them *acts* here to not confuse them with the research purposes mentioned earlier (i.e., developing an intervention, answering local questions, or answering questions that are important to the scientific community).

Several authors have put forward suggestions on what are the research acts in social science research. These include Gall, Borg, and Gall (1996), Jarvinen (2000), Mark, Henry, and Julnes (2000), Shadish, Cook, and Campbell (2002), Stokes (1997), and Yin (2003). However, I have found it helpful to use the following categories of research acts to describe the kinds of activities that researchers do and the kinds of research questions they ask. Those categories are *identification*, *description, comparison*, *correlation, experimentation*, and *explanation*.

In some sense these research acts, from identification to explanation, are linear in their degree of ability to explain causal mechanisms. For example, one has to identify a causal factor to be able to explain how it works in a causal model. However, that linearity does not mean that one type of research act necessarily needs to precede another type of research act. For example, one does not necessarily need to do experimentation or correlation to make a causal explanation. And, it does not mean that one cannot switch back and forth repeatedly between research acts. For example, in grounded theory research one iterates between cycles of identification, description, and explanation to arrive at a theory based on the data gathered.

Identification. The first research act, which could just have easily have been labelled exploration or orientation, deals with becoming aware of a phenomenon, its contexts, and its constructs. For example, in order to create a quantitative survey to measure the degree of users’ reactions to a new technological innovation, first one would have to identify the types of reactions that one wants to measure. Similarly, in qualitative research one might first have to establish that a phenomenon exists before describing the attributes or elements of that phenomenon. Identification is often the purpose of quantitative correlational research approaches and in many qualitative research approaches.

Description. One might use quantitative or qualitative description to describe the attributes of a phenomenon that came to light through the act of identification. In qualitative descriptive research, for example, if client satisfaction is identified as an important factor in some phenomenon, then a researcher might do a qualitative study to provide a detailed description of the attributes of client satisfaction. It might turn out that client satisfaction has several sub factors, and the researcher might have to revert to identification to become aware of those sub factors and, then, back to description to describe their attributes. In quantitative descriptive research, a researcher might give out a survey to measure the degree of satisfaction clients report for each of the sub factors identified earlier.

Research questions that relate to quantitative description often begin with questions phrases such as “How many...”, “What percentage of...”, “How often ...”, and so on. Research questions that relate to qualitative description begin with phrases such as “What kind of ...”, “What are the properties of ...”, “What is the meaning of ...”, “What are the types of ...”, and so on.

Comparison. The next type of research act, comparison, consists of two or more instances of description and an analysis of how those instances of description differ. In the field of educational technology, a researcher might examine how the gaming choices of male students differ from the gaming choices of female students. Research questions that relate to comparison involve differences – for example: “Do expert and novice programmers differ in the how they use algorithm animation software?”

In comparison studies, researchers do not manipulate variables and do not assign participants to treatment or control groups. The point of contrast in a comparison study is usually on some normally nonmanipulable attribute, such as age, mother tongue, gender, or previous experience, and so on. In health research, comparison studies (also called causal-comparative studies or case-control studies) are frequently seen because often it is not ethical or possible to assign people to treatment and control groups. Comparative studies are useful in those cases when the effect is known, but the cause is not known or cannot be manipulated (Shadish, Cook, & Campbell, 2002).

Correlation. Correlation consists of multiple instances of comparison to examine the (co)relationships between variables. For example, an educational technology researcher might be interested in knowing whether the use of a certain feature in a technological intervention is related with an increase in academic achievement. One practical outcome of examining correlations is that, under instance of high correlation, predictions about the behavior of one variable can be made from the behavior of correlated variables. Some examples of correlational research questions follow: “Is there a relationship between the number of hours a day spent watching educational programs and academic achievement?” or “Is there an association between the number of people in an online classroom and attrition?”

One important note is that correlation does not prove causation. Many occurrences are correlated but are not causally linked. *Confounding factors* can mask an actual association or make it appear that an association exists when one really does not. For example, there is a positive correlation between the sale of cooling fans and drowning deaths, but obviously, one does not cause the other. The confounding factor is that the heat of the summer months is correlated with both an increase in the sale of cooling fans and with the number of people who go swimming (and subsequently drown).

While correlation does not prove causation, it can be an initial clue that a causal relationship exists. The type of research act discussed next, experimentation, can be used to help determine if a causal relationship does indeed exist.

Experimentation (causal description). Shadish, Cook, and Campbell (2002) define an experiment as "a study in which an intervention is deliberately introduced to observe its effects" (p. 13). An educational technology researcher might conduct an experiment, for example, by introducing a new version of a technological tool and comparing the academic results between the phases when the students used the newer version of the tool and the phases when the students used the older version. Experimentation might be thought of as a special case of comparison in which the researcher changes something about a situation and then makes a comparison. Or it might also be thought of a special case of description in which "the consequences attributable to deliberately varying a treatment" (Shadish, Cook, & Campbell, 2002, p. 9) are described. Some examples of experimental research questions follow: "What are the effects of using a virtual data collection tool on the quantity of data that are collected?" or "Does technology intervention X or Y lead to better mathematics achievement?"

While experimental research is prized for its ability for the causal description of phenomena, there are a few important caveats about experimental research and causal claims that need to be mentioned. First, while experimental research can generate information that can help support causal claims, it dos not guarantee causal certainty. Experimentation is a means, not an end, to arriving at sound causal claims. Shadish, Cook, and Campbell (2002) stated this point well. They wrote,

> *Experiments yield hypothetical and fallible knowledge that is often dependent on context and imbued with many unstated theoretical assumptions. Consequently, experimental results are partly relative to those assumptions and contexts and might well change with new assumptions or contexts....to the extent that experiments reveal nature to us, it is through a very clouded windowpane. (p. 29)*

Second, experiments are good at causal description – that is, "in describing the consequences attributable to deliberately varying a treatment" (Shadish, Cook, & Campbell, 2002, p. 9) – but are not so good at causal explanation – that is, in "clarifying the mechanisms through which and the conditions under which [a] causal relationship holds" (p. 9). For example, by flicking a light switch on and off and observing the light going on and off, one could easily use causal descriptive reasoning to conclude that flicking the light switch causes the light to go on or off. But knowing *that* flicking the light switch causes the light to go on is much different than knowing *why* or *how* flicking the light switch causes the light to go on.

Causal explanation. As mentioned above, experimentation produces data that is useful for causal description. Unlike causal description though, which is used for determining that a certain cause leads to a certain effect, causal explanation can be used for explaining *why* or *how* a certain cause leads to a certain effect. Causal explanations often come about by examining a phenomenon in great detail.

Coming back to the light bulb example, if the goal were to provide an explanation for why turning on a light switch causes the light bulb to go on, a researcher using causal explanation would look into the walls and examine the wires, bulbs, switches, fuses, circuit breakers, and so on. From that, the researcher could come up with an explanation of how flicking a switch ultimately leads to light being emitted from a bulb. By doing pattern matching between what elements theoretically are needed to make a light bulb work and what elements are actually in place, the researcher could even determine that flicking the switch would turn on the light without ever having to actually flick the switch.

There have been many useful descriptions of how causal explanation works. Scriven (1976) describes causal explanation as a research act that uses a *modus operandi* approach – the same approach that a doctor uses to make a diagnosis or the same approach that a detective uses to catch a criminal. In short, in the modus operandi approach an observed pattern (e.g., a set of symptoms that a patient has) is matched with a known set of patterns (e.g., the set of symptoms associated with a particular illness). The often heard phrase in criminal investigation programs – this (pattern of evidence) is consistent with that (criminal phenomenon) – is evidence of the modus operandi/pattern matching approach in action. Mohr (1999) describes causal explanation as a research act that uses physical causal reasoning – the same reasoning that lets physicists predict the movement of objects. By knowing the theories that underlie physical causes, physicists can make causal explanations of physical phenomena. However, the theories of human behavior are much different than the theories of physical motion. Shadish, Cook, and Campbell (2002) describe causal explanation, not exclusively, as multiple cases of causal description. Whichever characterization of causal explanation one adopts, the essence is that it allows one to explain why or how causal systems work.

Dimensions in Research Acts

While research acts can be categorized as identification, description, comparison, correlation, experimentation, or explanation, it is also helpful to consider other dimensions – including whether the research adheres to qualitative or quantitative traditions and the degree to which the research is generalizable, accurate, feasible, appropriate, and useful.

General vs. local. One key dimension in research is to what degree results are local or general – that is, the degree to which results are generalizable across units, treatments, outcomes, or settings. In some cases it is sufficient to make local conclusions – that is, conclusions that are meant to be generalized only to local participants, treatments, outcomes, or settings. For example, in a program evaluation, it is probably sufficient to conduct research that only applies to the program being evaluated because the funders of the evaluation are primarily interested in the results of their program and not necessarily interested in the results of other programs. But, stakeholders in similar programs would probably be interested. In most cases in traditional education research, conclusions have more worth if they are generalizable – that is, if the conclusions apply to other units, treatments, outcomes, or settings outside of the original setting. In fact, Stanley and Campbell (as cited in Shadish, Cook, & Campbell, 2002, p. 97) argue that in research on teaching, “generalization to applied settings of known character is the desideratum.”

Qualitative vs. quantitative. Traditionally, some of the research acts described above have been associated with either qualitative or quantitative traditions. For example, case study research has traditionally been associated with qualitative research; experimental research has been traditionally associated with quantitative research. However, there is no reason that either quantitative or qualitative methods could not be used in any of the research acts. Theoretically, one could do an experiment in which only qualitative data were collected. Similarly, one could do a case study in which only quantitative data were collected, as Yin (2003) points out. There is growing support for combining qualitative and quantitative types of data to create a variety of types of evidence to support a claim (Creswell & Plano Clark, 2006; Johnson & Onwuegbuzie, 2004).

Exploration vs. confirmation. Another dimension of research is to what degree the goal is to explore a phenomenon or to confirm (or help disconfirm) a pre-existing hypothesis. In some types of research, like grounded theory, the researcher refrains from making a research hypothesis until the data begins to accumulate. In that type of research, the researcher might have an idea or a topic to explore but does not try to gather evidence for or against any particular proposition. One could say that the exploratory researcher wanders in a specific direction. This type of research is often considered to be useful when there is little or no understanding of a phenomenon or when a line of research gets stuck and new hypotheses need to be generated (Strauss & Corbin, 1990).

In many other types of research the goal is to build evidence to help confirm (or disconfirm) a claim. For example, in hypothesis testing one creates a testable, a priori hypothesis that is usually based on previous research or theory. In this type of research, one arrives at knowledge by positing a variety of hypotheses, testing the validity of those hypotheses, and eventually deciding on which hypothesis of many is the most likely. For example, a researcher might posit from theory or previous research that the method of instruction is more important than the medium of instruction in terms of student academic achievement. The researcher would then conduct an experiment in which evidence could be gathered that would either support or discredit that hypothesis.

Another type of confirmatory research is replication research. In replication research, one replicates another researcher's investigation to see if the same results generalize across units, treatments, outcomes, or settings. While replication research is not generally given as much value by the scientific community as research that creates new information, replication is nevertheless a cornerstone of science and provides an excellent opportunity for beginning researchers to hone their craft.

In reality, exploratory and confirmatory approaches intertwine. The act of trying to carry out confirmatory research usually brings about new hypotheses about a phenomenon. Exploratory research that keeps ending up at the same conclusion can help build strong evidence for, or against, a claim.

Level of participation. This dimension involves two aspects: (a) how closely researchers become involved in the phenomenon and setting they are studying and (b) how closely the participants in a study become involved in the research process. In some types of research, like ethnographic research, the researcher becomes a part of the community being investigated (see LeCompte & Schensul, 1999). In other types of research there is a strict line between the researcher and participant. Also, in some types of research; like participatory action research or participatory evaluation; the participants collaborate with a researcher or the researcher acts as a facilitator for participants who carry out the brunt of the research (see Reason &

Bradbury, 2001). In other types of research, the researcher is the only person who participates in the design, data collection, analysis, and reporting of research. There are many ongoing debates and discussions about the pros and cons of the different degrees of researcher and stakeholder participation, but they are too numerous to go into here.

Accuracy. In some cases, it is necessary for research to have much accuracy; that is, it must produce sound information that is (a) comprehensive, (b) technically adequate, and (c) with judgments that are logically aligned with the data collected (Joint Committee on Standards for Educational Evaluation, 1994). For example, it makes sense to have much accuracy when lives and well-being are at stake or when policies or programs are involved that affect many people or require large amounts of resources. Also, in some cases accuracy is expected as a matter of fact, such as in dissertations or in articles in prominent journals. In other cases, however, less accuracy is acceptable. For example, it would certainly be impractical to conduct a randomized group experiment in every cycle of a design-based research study. Likewise, it would be impractical to spend a large part of an organization's resources on answering a large set of evaluation questions with much accuracy. Instead, it might be better to focus on answering the most important evaluation questions with more accuracy and answering the less important evaluation questions with less accuracy. Accuracy and feasibility are often tradeoffs.

One aspect of accuracy is reliability – the degree to which measurements are consistent over time, situations, or raters. Having high reliability is important in some types of research and less important in others. For example, high reliability of measurements might be important when creating an instrument to predict success in a graduate program in educational technology, but it would be less important in the early stages of a design-based research study when several informal investigations are being conducted to gain insights into how an intervention can be improved.

Utility. Ideally, research should be "informative, timely, and influential" (Joint Committee on Standards for Evaluation, 1994, p. 4). But, the import given to utility can vary across research traditions. For example, in formative evaluation the goal is to create information that will be used to improve educational programs or policies. The stakeholders need to be able to easily understand and use the evaluation information. In other types of research, like basic research, the utility of the research is expected in the future. Although the utility of basic research is latent, basic research has been shown to be an essential factor in a large proportion of major breakthroughs (Comroe & Drips, 1976). Also, what kind of evidence is considered to be actionable or useful varies across settings and audiences. Remember that the crux of the methods-choice debate is deciding "when, or under what circumstances, various methodologies provide the most useful, or actionable, evidence" (Julnes & Rog, 2007a, p. 2).

Propriety. Propriety, the degree to which the rights of individuals involved in research are protected (Joint Committee on Standards for Educational Evaluation, 1994), is a critical dimension in all types of research. However, different types of research have different types of propriety issues. For example, meta-analytic research – research about research outcomes – generally does not involve propriety issues dealing with the treatment of human participants; however, it does involve other propriety issues involving complete and fair assessment, disclosure of findings, conflicts of interest, and possible fiscal responsibility. In other types of research, the ethical treatment of human participants is a critical factor in the choice of a methodology. For example, deciding on whether to use a randomized experiment or some other research design can hinge on the ethical issues involved. For example, Boruch (2007) puts forth a set of questions, which follow, to determine if a randomized experiment is ethically justifiable:

- *Is the social problem serious? If the answer is yes, then consider a randomized trial to evaluate the purported solutions. Otherwise a trial is not worthwhile or ethical.*
- *Are purported solutions to the problem debatable? If the answer is yes, then consider doing a randomized trial. If the answer is no, then adopt the purported solution.*
- *Will randomized trials yield more defensible (less equivocal and unbiased) results than alternative approaches to estimating effects? If the answer is yes, consider mounting a randomized trial. If the answer is no, then rely on the alternative approach.*
- *Will the results be used? If the answer is yes, then consider mounting a randomized trial. If not, forget about the trial, or redesign the randomized trial so that rights are protected. (pp. 56–57)*

Feasibility. Feasibility, the degree to which research does not consume more time, money, or resources than necessary, is also an important consideration in research design. Some research traditions, like design-based research, are based on repeated, rapid cycles of investigation. And, therefore, it would be impractical, if not impossible, to do a randomized trial each iteration. In short, one has to weigh the costs of each type of research design with the benefits that could potentially come about. Also, what may be feasible in one setting might not be feasible in another.

Conclusion

In summary, there is no simple answer for which method is most appropriate for a given situation. As discussed above, there are many factors that influence methods choice. There are factors that influence the formation of the research question: (a) the research problem, (b) the purposes of the research and their corresponding traditions, and (c) the state of the previous research. There are also factors that influence how a research question is answered: (d) the research act implied in the questions, (e) the feasibility of the research, (f) safeguards for propriety, (g) the degree of utility needed, (h) the degree of generalizability needed, (i) the degree of stakeholder participation in the research, and (j) the degree of researcher participation in the research setting, among other factors. For each research situation, the researcher must carefully weigh these factors together to finally decide on which research methods to use.

To aid in this process of considering and weighing the methodological factors mentioned here, a list of key questions in educational technology methods choice can be found in the Appendix to this book. With hope, this list of questions will be useful for helping student researchers think through the considerations involved in methods choice and as an instructional aid for those who teach or supervise students of educational technology.

Questions to Consider

1. What are the similarities and differences between design-based research, evaluation research, and education research? In what ways can they be used to inform each other?

2. In this chapter I reported on the types of educational technology research questions that were often asked in the past. What types of educational technology research questions do you think will be asked in the future?

3. Collect a few educational technology research questions. What are the research acts implied in the questions?

4. What is your response to my claim that quantitative or qualitative data can be useful in any of the research acts mentioned above?

5. What do you think are the advantages and disadvantages of a researcher's being a participant and participants' being researchers?

Chapter 2

Types of Research Approaches in Educational Technology Research and Development

In the previous chapter, I discussed the factors that influence methods choice in educational technology research. In this chapter, I describe the major types of research approaches that are used in educational technology research and how they relate to the research acts mentioned earlier.

Although there are many ways to categorize research approaches, I have chosen to use an adaptation of the categories presented in Gall, Borg, and Gall (1996) for the quantitative research approaches and Creswell (2007) for the qualitative research approaches.

The quantitative approaches that I discuss are:

- Survey research,
- Causal-comparative research,
- Correlational research, and
- Experimental research.

Also in the section on experimental research I discuss the major threats to internal validity and the basic categories of experimental research designs.

Concerning the types of research approaches typically associated with qualitative data, I discuss the five following major qualitative traditions:

- Narrative research,
- Phenomenological research,
- Ethnographic research,

- Case study research, and
- Grounded theory research.

Although some of these research approaches are used less often in educational technology research than others, I, nonetheless, have included the rarely seen approaches (e.g., narrative research and ethnographic research) here because of their high potential for answering important questions in educational technology research and development.

Survey Research

Survey research, whose related research act is attribute description, typically answers questions that deal with the frequencies or proportions in a population. An election poll is an example of survey research. Survey researchers might, for example, be interested in the percentage of people who will vote for a given candidate or the percentage of people who reported being satisfied with a piece of educational software. While questionnaires are common quantitative descriptive research instruments, any type of observational method – from standardized tests to direct observation – can be used to determine frequencies or proportions within a population.

In general, survey researchers either conduct surveys or censuses. In surveys, data are only collected from a representative sample and the results are inferred to the whole population. Surveys are conducted because it is often impractical or impossible to collect data about an entire population. For example, in an election poll, survey researchers might interview a representative set of individuals about how they are going to vote and, if the set is truly representative, the researchers can make a good prediction about how the population is going to vote. In censuses, survey researchers collect information about every individual in the population. For example, the U.S. government conducts a census in which they strive to gather information about every individual living in the United States. However, because of the time and expense involved they only conduct a census once every ten years.

To make a valid inference from a sample to a population, it is critical that the sample be representative of the population. One method that can help achieve the representativeness of the sample is through random sampling – by choosing cases randomly from the population. Random sampling, which is also referred to as *random selection,* is also a requirement of many statistical tests. Random sampling, however, does not ensure that a sample is representative. It is unlikely, but possible, that a random sample will not constitute a good representation of the population. The more cases that are randomly sampled, in a statistical sense, the more likely it is that the percentages in the sample are near the percentages in the population.

Often it is impractical or impossible to take a random sample so, instead, researchers take a *purposive sample of typical cases* – sampling in which the goal is to purposefully choose a set of cases that is representative of the population. Basically, one can make an argument for the representativeness of a purposive sample, or any other type of sample, by assessing in what ways the sample is similar to the population and in what ways it is different (Shadish, Cook, & Campbell, 2002). For example, if a researcher wanted to make an argument that the students in his or her first-year undergraduate computer science class are representative of students in first-year undergraduate computer science classes in general, then that researcher might collect information about the proportion of males to females, the students' previous experience in computing, their socio-economic status, and so forth, and compare that information with information that is known about the population. If it turned out that the students in the class were all female Harvard students who went to a computer-science-oriented high school, then it might be hard to make an argument that those students are representative of first-year undergraduate computer science students in general.

Besides purposefully sampling typical cases, a researcher might also use other types of purposive samples. For example, a researcher might be interested in extreme cases and purposefully sample the cases that exhibit extreme qualities.

Causal-Comparative Research

In causal-comparative research, which is also called *case-control* research, one typically compares a group to one or more different groups or compares the same group at different times and *does not manipulate a variable*. For example, a researcher might examine whether male students differ from female students in how they experience community in online courses or the researcher might examine how perceptions of online learning have changed over the years.

Shadish, Cook, and Campbell (2002) argue that causal-comparative research is useful in situations in which an effect is known, but the cause is not known. For example, causal-comparative research might be used to determine what caused students to drop out of an educational program by determining how those who dropped out and those who stayed in the course differed. Causal-comparative research is also useful in those situations where the cause (the independent variable) cannot be manipulated. For example, it is not realistically possible to manipulate a participant's gender or background. In these situations, comparative research can be used to gather evidence for, or against, a cause-effect relationship between variables; however, it is difficult to establish this causal relationship with any certainty using comparative research.

One problem is that in causal-comparative research it is difficult to rule out all causes other than the one being investigated. For example, imagine that a researcher wants to find out if online learning causes more academic achievement than traditional classroom teaching and suppose that it turns out that students in traditional classrooms have better academic achievement than students in online classrooms. In this case, it would be very difficult to establish that traditional classroom learning caused the better academic achievement because there are so many other factors that could have caused the increased achievement. For example, perhaps it is the case that online learners typically have full-time jobs and cannot devote as much time to study as students who can attend traditional classroom lectures. So, the real cause of the difference in academic achievement could actually be available study time. The crux of the matter is that in causal-comparative research it is extremely difficult to substantiate causal claims without a great deal of lateral research that systematically rules out confounding causal factors.

In longitudinal research, which is appropriate for asking questions about developments over time, there are four types of studies that are often used. In *trend studies* a researcher chooses the same theoretical sample, but measures the theoretical sample at different times and with different participants. For example, in a trend study a researcher might choose 30 students from the incoming group of first-year computer science students and repeat this every year – e.g., 30 students from all of the computer science students beginning in 2000, then 30 students from all the computer science students beginning in 2001, then 30 students from all the computer science students beginning in 2002, and so on. In this way the researcher could determine if there is a change in the previous computing experience of first-year computer science students over time.

Another type of longitudinal study is the *cohort study* in which a researcher chooses a different sample at different points in time and the population remains constant. For example, a researcher might choose a different set of 30 students, each year, from the cohort of students who began in 2000. To illustrate, in 2000 the researcher would sample thirty students from the 2000 cohort. In 2001, the researcher would choose another sample of students from the 2000 cohort and in 2002 the research would choose yet another sample of students from the 2000 cohort.

The third type of longitudinal study is a panel study in which a researcher chooses the same set of cases every year. For example, the same set of 30 students who originally had been chosen would be the same set of students who would be chosen every subsequent year.

The final type of longitudinal study is a cross-sectional study in which data are obtained at only point in time, but the cases are chosen from different age groups. For example, in a cross-sectional study one might choose 30 students who started in 2001, 30 students who started 2002, and 30 students who started in 2003 to determine if there is a difference in the cohorts of incoming students over time.

Each of these types of longitudinal research has its advantages and disadvantages. While panel studies are more sensitive to small changes over time than are trend, cohort, or cross-sectional studies, they are more difficult to carry out and there are problems with attrition (i.e., problems with people dropping out of the study) and problems with participants being influenced by repeated measurements. On the other hand, trend cohort, and cross-sectional studies are easier to carry out and do not suffer from attrition or repeated measurement problems; however, they do not lead to data as rich or valid as panel studies.

Correlational Research

In correlational research, the researcher is interested in how one or more variables change in relation to how other variables change. For example, a correlational researcher might be interested in whether academic achievement increases, stays the same, or decreases as the amount of dialogue in an online course increases. Besides examining the relationships between two variables, correlational research can be used for prediction or for examining the relationship between two or more variables with the influence of other variables factored out (e.g., regression analysis; see Cohen & Cohen, 1983), for quantitatively creating or confirming complex causal models (e.g., exploratory and confirmatory factor analysis; see Thompson, 2004), or for classifying data into groups (e.g., cluster analysis; see Romesburg, 2004).

Correlational datasets are often represented in a scatter plot in which data points are plotted on two dimensions. Figures 4, 5, and 6 are examples of scatter plots on which the values on the horizontal axis (x) and the values on the vertical axis (y) are charted together. If two variables increase or decrease together they are said to be *positively correlated* (see Figure 4). If one variable increases while the other dec-reases, the variables are said to be *negatively correlated* (see Figure 5). And, if the variables tend to be independent, that is if an increase or decrease in one variable in not accompanied by an increase or decrease in another variable, then the variables are said to have no correlation (see Figure 6).

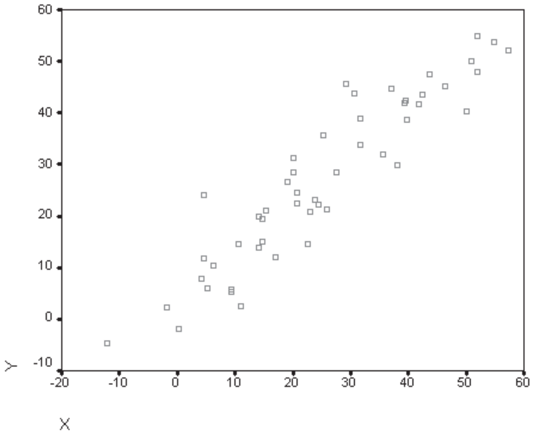


Figure 4. *Scatter plot of a strong positive correlation, (r = .93).*

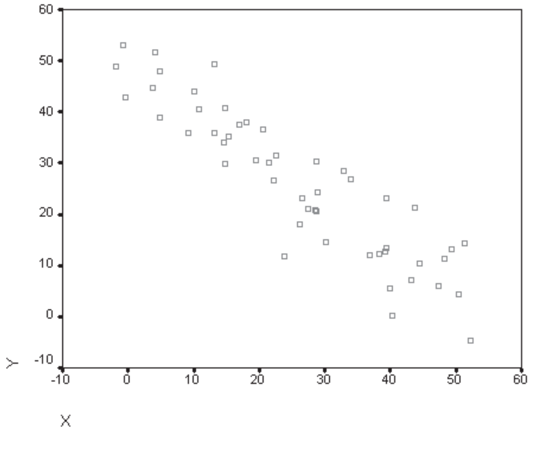


Figure 5. *Scatter plot of a strong negative correlation, (r = -.91).*

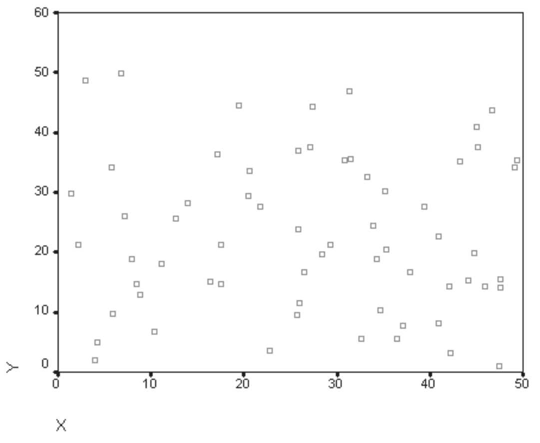


Figure 6. *Scatter plot when there is practically no correlation, (r = -.02).*

Although this book is not meant to be a treatise on statistical practices in the field educational technology, it might be important to note that the most frequently used measure of correlation is the Pearson Product Moment Correlation, which is denoted by an *r,* and its values range from 1.0 to -1.0. A positive value of *r* indicates that there is a positive correlation; a negative value indicates that there is a negative correlation, and a value of zero indicates that there is no correlation. Whether the correlation is positive or negative shows the *direction* of the correlation. The closer a correlation is to 1.0 or -1.0, the closer it is to being perfectly correlated. When there is a perfect correlation it is possible to know what the value of one variable will be by knowing the value of another. The nearness to 1.0 or -1.0 shows the *degree* of the correlation.

Figures 4, 5, and 6 have an *r* of .93, -.91, and -.02, respectively. So, the scatter plot in Figure 4 shows a very strong correlation in the positive direction, the scatter plot in Figure 5 shows a very strong correlation in the negative direction, and the scatterplot in Figure 6 shows that there is practically no correlation. Since there is practically no correlation in Figure 6, the direction is largely irrelevant.

One note about correlational research, which also applies to comparative research, is that it is very difficult to gather enough evidence to have a high degree of certainty that a causal relationship exists. Correlation does not prove causation. For example, there might be a correlation between variable *a* and variable *b* because variable *a* causes a change in variable *b*; because variable *b* causes a change variable *a*; or that a third variable, say variable *c*, is correlated with and causes a change in both variable *a* and variable *b*. On this point, Gall, Borg, and Gall (1996) wrote,

> *Correlational coefficients are best used to measure the degree and direction (i.e., positive or negative) of the relationship between two or more variables and to explore possible causal factors. If a significant relationship between variables is found, their causality can be tested more definitely by using [other methods, like the experimental method]. (p. 414)*

In the following section, I discuss the experimental method for arriving at causality – by what Francis Bacon called "twisting the lion's tale" (as cited in Shadish, Cook, & Campbell, 2002, p. 2).

Experimental Research

Experimental research is a subset of comparative research in which the researcher manipulates a variable (called the *independent variable)* and determines how changing that variable effects one or more outcome variables (called *dependent variables*). For example, an experimental researcher might implement a new technological tool (the independent variable) in a classroom and then see how the implementation of that tool effects student academic achievement (the dependent variable). The crux of experimental research is that a researcher compares a factual condition (i.e., what happens when the variable is not manipulated – the control or baseline condition) with one or more counterfactual conditions (i.e., what happens when the variable is manipulated – the experimental condition) so that the unique effect of the variable's manipulation can be measured. An experimental researcher carefully chooses research designs to lessen the threat that something other than the variable being manipulated causes the outcome that is measured. Fortunately, much research has been conducted on these threats that can lead to erroneous causal conclusions. These types of threats are called *threats to internal validity* and are the subject of the following section. Understanding these threats helps makes it possible to understand the rationale for the different types of experimental research designs.

Major Threats to Internal Validity

To give meaning to experimental research designs, it is necessary to understand the major threats to internal validity. Remember that experimental research is particularly suited to determining causal relationships, but threats to internal validity are factors that can make it appear that the independent variable is causing a result when in fact it is a different variable or set of variables that is causing the result. In other words, threats to internal validity in experimental research are *hidden causes*. I use an acronym, HARM-ITS, as a mnemonic device to remember the seven major threats to internal validity: *history, attrition, regression to the mean, maturation, instrumentation, testing*, and *selection*.

History. Sometimes an unintended event outside of the experiment can affect the outcomes of an experiment. For example, suppose that there is loud construction work going on in a building where students are taking a test. It might turn out that the construction noise, and not a poorly designed intervention, could be the cause of low test performance.

Attrition. Participants have a tendency to drop out of experiments over time. Sometime the drop out is random and sometimes the drop out is non-random. When the drop out is connected with some factor that could affect the results, then attrition becomes a threat to internal validity. Imagine that there are 30 high-achieving students and 30 low-achieving students in an experiment. Suppose that 20 of the 30 low-achieving students dropped out of the experiment because the intervention seemed too difficult in the beginning and all of the high-achieving students stayed in the experiment. In this case, the final results would be biased toward the high-achieving students and the outcomes would appear better than they actually would have been if the low-achieving students had stayed in the experiment.

Regression to the mean. To the degree that a measurement is influenced by chance, there is a tendency for an individual who has an extremely high or extremely low score on a first measurement to have a score that is nearer to the mean on the next measurement. This phenomenon is called *regression to the mean.* Regression to the mean becomes a problem when individuals are selected into an experiment on the basis of a selection test *and* when that selection test is also treated as a pretest. For example, if a researcher were interested in only investigating high-aptitude students, then he or she might give out an aptitude test (i.e., a selection test) to a diverse set of students to determine which students have high-aptitude and, therefore, are eligible to be selected into the study. Regression to the mean would be a threat if the researcher treated the results of the selection test as if they were the results of a pretest. In this case, for example, if some of the selected students had, by chance, gotten atypically high scores on the joint selection/pretest, their scores on a subsequent measure probably would be nearer to their typical scores (i.e., scores nearer to the mean) just because of regression to the mean. This is because it is statistically unlikely to get atypically high scores just by chance, two times in a row. In this case, regression to the mean would make the intervention look less effective then it actually might be. The way to avoid regression to the mean is to keep the selection and pretest measurements separate. A researcher should give a selection test, select students, *then* give a separate pretest and posttest.

Maturation. Participants have a tendency to mature and develop over time, regardless of an intervention. For example, it is a fact that as young students progress through puberty and onto adulthood their social skills typically develop. So in an experiment dealing with an intervention designed to increase the social skills of teenagers, then maturation rather than the intervention might be the cause of increased social skills over time.

Instrumentation. When measurement instruments, including human raters, change over time, then the experimental outcome could be the result of the change of the instrument rather than the intervention. Say for example that a pretest turns out to be much harder than a posttest. In that case, a totally ineffective intervention might appear to have been effective, when in reality the difference in difficulties of the two tests was what caused the apparent change between pretest and posttest measurements.

Testing. Sometimes simply the act of taking a test, or being measured, can affect the outcomes of that test or measurement. Organizations who train individuals to take tests capitalize on this fact by having their clients take a test repeatedly. The act of repeatedly taking a test will often cause the results of a test to improve with each repetition.

Selection. The threat of selection occurs when the participants in the experimental and control conditions are not equal to begin with. Say for example that a group of researchers conduct an experiment in which they introduce an educational intervention in one classroom, do not introduce the intervention in a different classroom, and then measure academic achievement at the end of the year in both classrooms. It might be entirely possible, even probable, that one class had better performance at the beginning of the year and, therefore, that class would be expected to have better scores at the end of the year, regardless of the intervention.

One way to reduce the threat of selection is to randomly assign participants to either control or experimental conditions. Strictly speaking, in experimental research in which there are two or more groups, if there is not random assignment then that type of investigation is called a *quasi-experiment*, not an *experiment*.

There has been an ongoing debate in the education research community about the relative values of quasi-experimental and experimental research (Julnes & Rog, 2007a). Some groups, like the U.S. Department of Education, give funding priority to randomized experiments, which are sometimes considered to be the "gold standard" of research designs. However, in the field of education and educational technology, it is a rare occasion indeed when a true experiment can be conducted. In education, participants often are entitled to self-select their interventions and it is often difficult to ethically justify assigning students to experimental and control groups by chance, especially when the intervention is thought to be effective. Therefore, other groups like the American Education Research Association and the American Evaluation Association argue that carefully conducted quasi-experimental research can lead to results as valid as the results in experimental research. See Julnes and Rog (2007a) for a thorough discussion on this issue.

One way to avoid the threat of selection is to use group experimental designs that include a pretest. Even if there are differences on the pretest, in some cases, statistical techniques (such as ANCOVA; see Stevens, 1999) can be used to artificially adjust for the differences in pretest scores.

Group Experimental Designs

There are two basic dimensions to group experimental research designs. There is a within-subjects dimension and a between-subjects dimension. In the within-subjects dimension, the same participants (or participants who are matched) are measured at different points in time. For example, a group of participants might be given a pretest measurement and a posttest measurement. Each participant's pretest data would be compared with his or her own posttest data to determine if the intervention had been effective. In the between-subjects dimension, different groups of participants are measured. For example, a group that received an intervention might be compared with another group who did not receive the intervention. The researcher would then compare the outcomes of one group to the outcomes of the other group to determine if the intervention had been effective. With these two dimensions in mind, it is then possible to make sense out of most of the basic experimental research designs. As I will explain below, some experimental research designs are only within-subjects, some are only between-subjects, and some are a combination of both. The experimental designs that I deal with here are discussed in more detail in Shadish, Cook, and Campbell (2002).

The one-group posttest-only design. As the name implies, in this research design there is only one group who is given a test after an intervention has been delivered.

This design is subject to almost all threats to internal validity, because many factors besides the intervention could have caused the outcome. The one-group posttest-only design is only recommended when there is considerable prior knowledge about the dependent variable, so that a comparison can be made between the outcomes of the experiment and the outcomes predicted by previous research or theory (Shadish, Cook, & Campbell, 2002). As I will show in the following chapter, the one-group posttest-only design is common in educational technology research.

One popular, but questionable, variation of this design is to use a retrospective posttest in which participants try to judge how much they learned as a result of the intervention. These types of retrospective reports are however known to be quite biased (Silka, 1989). These retrospective tests might also be thought of as survey research about participants' self-reports of learning.

The posttest-only design with non-equivalent control groups. In this between-subjects design, the researcher compares the posttest results of an experimental group to the results of a control group. Adding a control group significantly reduces most threats to internal validity. For example, if a history threat were to happen it would probably happen to both groups equally and, in effect, would cancel itself out. Although this design lessens many threats, it is still vulnerable to the threat of selection. The groups may differ initially before the intervention and, therefore, would probably differ even after an ineffective intervention. Adding a pretest measurement is one way to improve on the posttest-only design with non-equivalent control groups.

The one-group pretest-posttest design. This is a within-subject design where one group is given a pretest measurement, an intervention is introduced, then a posttest measurement is given. In this design, selection is not a threat because there is not a control group; each participant acts as his or her own control. History, attrition, testing, regression to the mean, instrumentation, and maturation are the major threats in this design. It can be improved upon by adding a control group.

The untreated control group design with dependent pretest and posttest samples. This design has both within-subjects and between-subjects dimensions. There is a control group and experimental group and there are measures at two or more points in time. Within the groups, the pretest data for each participant is matched with his or her own posttest data. The major advantage of this design is that it combines features of within-subject and between-subject designs and, therefore, greatly reduces threats to internal validity. Also, in some cases it is possible to statistically control for differences between the experimental and control groups.

The list of experimental research designs mentioned above is by no means complete; however, most of the other designs are variations on the ones above. Shadish, Cook, and Campbell (2002) is an excellent resource for those are interested in learning more about experimental research design.

Single-Participant Research

Single-participant research is a type of experimental research in which the emphasis is on the performance of an individual over time rather than on group performance. In single-participant research, the researcher takes several measurements of one case over time. Over those time periods the researcher will alternately implement and withdraw an intervention (or variations of an intervention) and see how the outcome measurements change. For example, a researcher might measure a student's off-task behavior for five days to establish a baseline, implement an intervention and measure the results for five days, and then withdraw the intervention and see if the results return to the baseline. This type of single-participant design is called an ABA design, where A is a time period when there is no intervention and where B is a time period where there is an intervention. Alternately, a researcher could have a baseline period, introduce the intervention, introduce a variation of the intervention, reintroduce the original intervention again, then withdraw the intervention altogether. This would be classified as an ABCBA design, where C is the time period where the variation of the intervention is introduced.

Typically, single-participant data are analyzed visually. Figure 7 is a graph of hypothetical results from a single-participant ABA design, in which there are five measurements in each phase. Suppose that the result being measured is percent of intervals observed in which a student was exhibiting off-task behavior (e.g., disrupting others). In the initial A phase, the intervals of off-task behavior were high. In the B phase when an intervention was implemented, the intervals of off-task behavior decreased. And, in the final A phase when the intervention was withdrawn, the intervals of off-task behavior decreased again. This high-low-high pattern provides evidence that the intervention worked in decreasing off-task behavior. Single-participant graphs are also useful because they intuitively illustrate variation between measurements and trends over time. Kazdin (1982) is a classic and highly recommended text on single-participant methods.

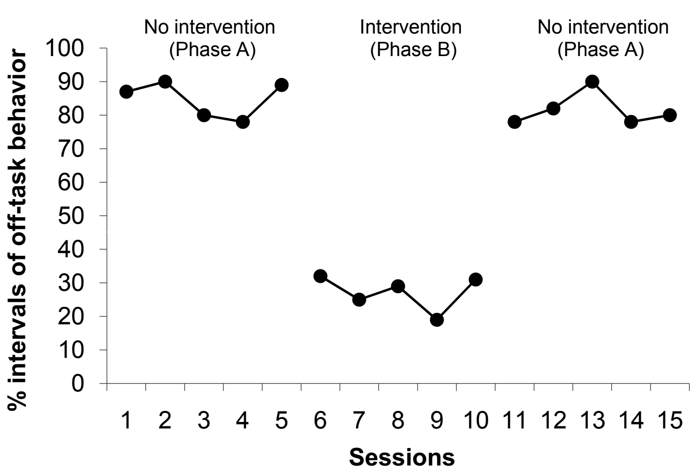


Figure 7. *Example of a single-participant graph.*

Single-participant research is well suited for design-based research. First, it does not require many participants, as group experimental research does, but it can generate strong evidence about causal relationships nonetheless. Second, single-participant research is flexible. The researcher can make impromptu variations to the intervention to quickly identify which variations of an intervention work and which do not. The researcher can also make impromptu variations to the environment to see which environmental factors affect the results of the intervention and which do not. One downside though to single-participant research is that it only works in situations where testing is not a threat to validity, since many measurements need to be made. Despite the suitability of single-participant studies in design-based research, it is surprising that single-participant studies are not used more frequently in design-based educational technology research (as I show in the next chapter).

The Five Major Qualitative Approaches

Coming back to the research acts discussed earlier, qualitative methods are ideal for creating thick, rich description (Geertz, 1973). Also, unlike quantitative descriptive research; which involves answering *who, what, where, how many*, or *how much* questions; qualitative research often involves *how* or *why* questions.

Through thick, rich description qualitative methods can be used alone or paired with other research approaches to help confirm causal conclusions (Mohr, 1999; Yin, 2003). To give an example of pairing research methods, a qualitative study might be initially used to create a theory about a phenomenon and identify the important variables, experimental research could be used to determine if that theory holds up and is generalizable. Or alternately, qualitative research could follow experimental research to help confirm that a causal relationship holds up when examined in detail.

In this section, I discuss the five major approaches to qualitative inquiry, as framed by Creswell (2007). They are *narrative research, phenomenological research, ethnographic research, case study research*, and *grounded theory research*.

Narrative Research

Narrative research is suited for situations when it is meaningful to tell the life stories of individuals. From these stories, insights about social phenomena and social meanings can be drawn. A narrative researcher typically collects data about an individual's life through interviews and documents. The researcher then retells the story and draws meaning and insights from that individual's life events. In educational technology research one might tell the life story of a teacher and focus on how technology has affected his or her classroom life. Recommended readings on narrative research include Clandinin and Connely (2000), Czarniawska (2004), and Denzin (1989).

Phenomenological Research

The goal of phenomenological research is uncovering the essence of a lived experience or phenomenon. Phenomenological researchers primarily gather data through interviews with individuals who have experienced a certain type of phenomenon. After bracketing (i.e., revealing) their own experiences about the phenomenon, the researcher picks out meaningful statements from the interview texts, gives meaning to those statements, and then goes on to provide a rich description of the essence of the phenomenon. In educational technology, for example, one might interview people who have participated in online courses and try to gain an understanding of how those people experience the phenomenon of *community* in online courses. Recommended readings on phenomenological research include Moustakas (1994) and van Manen (1990).

Ethnographic Research

In ethnographic research, the ethnographer strives to describe how a culture-sharing group works and the principles and values that under gird the group's behavior. Ethnographers tend to spend much time in the field making observations and interviewing the members of the culture-sharing group. In the field of educational technology, one might consider people who regularly participate in role-playing games as a culture-sharing group and, after collecting much detailed information, write a thick and rich description of how the culture-sharing group works and about the values that are shared. Recommended readings on ethnographic research include Atkinson, Coffey, and Delamont (2003), LeCompte and Schensul (1999), and Wolcott (1994, 1999).

Case Study Research

According to Yin (2003), "a case study is an empirical inquiry that investigates a contemporary phenomenon within its real life context, especially when the boundaries between phenomenon and context are not clearly evident" (p. 13). The goals of a case study are to develop an in-depth understanding of a case, or multiple cases, and to gain insight into the interaction between the phenomenon and the case. A case study researcher uses several sources of evidence; such as archival records, direct observation, interviews, and documents; and then analyzes that data though pattern-matching, explanation-building, and addressing rival explanations (Yin, 2003). Recommended readings on case study research include Merriam (2001), Stake (1995), and Yin (2003).

Grounded Theory Research

The goal of grounded theory research is to create a theory that is based on the data collected. According to Creswell (2007) this type of research is appropriate when there is no existing theory or when the existing theories are incomplete. Research questions in grounded theory research concentrate on how a phenomenon was experienced and how the process unfolded. Typically a grounded theory researcher conducts between 20 and 60 interviews, depending on how long it takes to reach a point of saturation – a point when no new information is gained from collecting more data. The researcher then analyzes these data by first looking for themes (open coding), assembles the data in new ways (axial coding), creates a story line that connects themes, or suggests new hypotheses. The ideal result of a grounded theory study is a substantive theory that explains the phenomenon. Recommended readings on grounded theory research included Charmaz (2006) and Strauss and Corbin (1990).

Questions to Consider

1. How do each of the research approaches presented in this chapter relate to the research acts present in Chapter 1?

2. Would you call a posttest questionnaire in which students are asked to retrospectively judge how much they learned from an educational intervention an example of survey research or an example of one-group posttest-only experimental research?

3. Why is it difficult to establish causality with causal-comparative or correlational research?

4. Why does Francis Bacon describe experimental research as "twisting the lion's tale"?

5. Which of the research approaches above do you most often associate with educational technology research and development? Why?

6. Do you agree with the U.S. Department of Education's judgment that randomized experiments are the gold standard of research? Or do you agree with others that research approaches other than randomized experimental research, such as qualitative research or quasi-experimental research, can lead to equally suitable evidence?

Chapter 3

A Meta-Synthesis of Methodological Reviews of Educational Technology Research

Thus far I have discussed the theoretical dimensions that can be used to chart educational technology research. In this chapter, I describe the research methods that educational technology researchers actually use in practice. I also compare those methods with the methods used in traditional education research.

This chapter is intended to serve two purposes. The first is to identify the methods that have been used in the past to inform the research of the present and future. My intention is to identify the strengths and weaknesses of the past research so that the strengths can be built upon and the weaknesses can be remedied. The second is to yield information that can advance the dialogue on educational technology as a discipline.

This chapter is organized around the following questions:

1. What are the meta-categories that can be used to subsume the research categories in other methodological reviews of educational technology research?
2. What are the proportions of quantitative, qualitative, and mixed methods research that educational technology researchers have tended to use?
3. How do those proportions differ over time periods and publication forums?
4. How do those proportions compare to the proportions in the field of education research proper?
5. In what proportions do educational technology researchers choose (a) research approaches, (b) experimental research designs, and (c) measures?
6. How do educational technology researchers tend to report educational technology studies?
7. What suggestions are given for improving educational technology research?

Since there have been many instances of research about research (meta-research) in educational technology already published, in this chapter I will simply concentrate on synthesizing that research. To answer Questions 1, 2, and 3, I will report on a review (Randolph, 2007a) in which I synthesized 13 methodological reviews of the educational technology research. For the answer to Question 4, I rely on Randolph (2007a) and a methodological review of education research proper by Gorad and Taylor (2004). (By *education research proper* I refer to the education research literature published outside of forums typically associated with educational technology.) Finally, to answer Questions 5, 6, and 7, I rely on a review of educational technology articles published in the proceedings of the International Conference on Advanced Learning Technologies (ICALT) (Randolph et al., 2005). In that the review, the authors reviewed all (126) of the full papers in the proceedings of ICALT 2004 – one of the premier IEEE-sponsored conferences in educational technology.

To answer Questions 5, 6, and 7, I generalize the articles in the proceedings of the ICALT conference to educational technology articles in general. My rationale for making this generalization is that the key variables in the ICALT review match the patterns of other broader reviews within and across fields, as I demonstrate later.

What are the Meta-Categories that Can be Used to Subsume the Research Categories in other Methodological Reviews of Educational Technology Research?

Methodological reviews, literature reviews that concentrate on research practice, have been published often in educational technology. Randolph (2007a) conducted an electronic search, hand search, and query to the members of ITFORUM to identify the complete set of those methodological reviews of educational technology. The criteria for inclusion and exclusion are listed below, (from Randolph, 2007a):

- *It was a quantitative review (e.g., a content analysis) of research practices, not a literature review in general or a meta-analysis, which focuses on research outcomes.*
- *The review dealt with the field of educational technology or distance education.*
- *The review was written in English.*
- *The number of articles that were reviewed was specified.*
- *The candidate review's categories were able to be subsumed under metacategories.*
- *The review's articles did not overlap with another review's articles. (When reviews overlapped, only the most comprehensive review was taken.) (p. 21)*

Thirteen methodological reviews of the educational technology literature were identified: Alexander & Hedberg, 1994; Caffarella, 1999; Clark & Snow, 1975; Dick & Dick, 1989; Driscoll & Dick, 1999; Higgins, Sullivan, Harper-Marinick, & Lopez, 1989; Klein, 1997; Phipps & Merisotis, 1999; Randolph, 2007c; Randolph, Bednarik, Silander, et al., 2005; Reeves, 1995; Ross & Morrison, 2004; and Williamson, Nodder, & Baker, 2001. Table 6 shows the research questions that were asked in each of the identified reviews. Five of those reviews (Alexander & Hedberg, 1994; Caffarella, 1999; Driscoll & Dick, 1994; Phipps & Merisotis, 1994; and Ross & Morrison, 2004) were ultimately excluded because they did not meet all of the criteria for inclusion listed above (see Randolph, 2007a, for an explanation). Table 6 shows the research questions that were addressed in each of the previous methodological reviews.

Table 6. Research Questions in Past Educational Technology Methodological Reviews.

Review	Overview of research questions
Alexander & Hedberg, 1994	What, and in what proportions, are evaluation models used in evaluations of educational technology?
Caffarella, 1999	How have the themes and research methods of educational technology dissertations changed over the past 22 years?
Clark & Snow, 1975	What research designs are being reported in educational technology journals? In what proportions?
Dick & Dick, 1989	How do the demographics, first authors, and substance of articles in two certain educational technology journals differ?
Driscoll &Dick, 1999	What types of inquiry are being reported in educational technology journals? In what proportions?
Klein, 1997	What types of articles and what topics are being published in a certain educational technology journal? In what proportions?
Higgins et al., 1999	What do members of a certain educational technology journal want to read?
Phipps & Merisotis, 1999	What are the methodological characteristics of studies published in major educational technology forums?
Randolph, 2007c	Are the same methodological deficiencies reported in Phipps & Merisotis (1999) still present in current research?
Randolph et al., 2005	What are the methodological properties of articles in the proceedings of ICALT 2004?
Ross & Morrison, 2004	What are proportions of experimental designs being used in educational technology research?
Reeves, 1995	What types of methodological orientations do published educational technology articles take? In what proportions?
Williamson et al., 2001	What types of research methods and pedagogical strategies are being reported in educational technology forums?

Note. From *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000-2005* (p. 24), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.

Table 7 shows the forums, the timeframe, and the number of articles reviewed in each review article. As Table 7 shows, the Randolph (2007a) meta-review included reviews of over 905 educational articles, from eight reviews over the past 30 or so years. The forums that were covered in the previous reviews were *AV Communication Review* (AVCR), *Educational Communication and Technology Journal* (ECTJ), *Journal of Instructional Development* (JID), *Journal of Computer-Based Instruction* (JCBI), *Educational Technology Research & Development* (ETR&D), *American Journal of Distance Education* (AJDE), *Distance Education* (DE), *Journal of Distance Education* (JDE), and *The Proceedings of the International Conference on Advanced Learning Technologies* (ICALT). The *mixed* forum articles were all from one review (Williams et al, 2001). Of the 46 articles reviewed in William et al.

> *37 originate[d] from refereed journals or conference proceedings and the remainder from academic websites or Government departments.... In particular we drew material from the conferences of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE) and from the National Advisory Committee for Computing Qualifications (NACCQ). (p. 568)*

Table 7. *Characteristics of Educational Technology Reviews Included in the Quantitative Synthesis.*

Review	Forum	Years covered	Number of articles reviewed
Clark & Snow, 1975	AVCR	1970-1975	111
Dick & Dick, 1989	ECTJ	1982-1986	106
	JID	1982-1986	88
Higgins et al., 1989	ECTJ	1986-1988	40
	JID	1986-1988	50
Reeves, 1995	JCBI	1989-1994	123
Klein, 1997	TR&D	1989-1997	100
Williamson et al., 2001	Mixed	1996-2001	46
Randolph, 2007c	AJDE	2002	12
	DE	2002	14
	JDE	2002-2003	40
Randolph et al., 2005	ICALT	2004	175[a]
Total			905

Note. AVCR = *Audio Visual Communication Review,* ECTJ = *Educational Communication and Technology Journal,* JID = *Journal of Instructional Development*, JCBI = *Journal of Computer-Based Instruction*, ETRD = *Educational Technology Research & Development,* AJDE = *American Journal of Distance Education*, DE = *Distance Education*, JDE = *Journal of Distance Education*, ICALT = *International Conference on Advanced Learning Technologies*. [a] 175 investigations reported in 123 articles. From *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005* (p. 25), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.

Across each of the reviews included in Table 7, Randolph (2007a) grouped the research categories used in each review into three meta-categories. The three meta-categories that emerged were *quantitative, qualitative, mixed-methods*, and *other*. The *other* category consisted mostly of papers that did not deal with human participants, such as literature reviews, program descriptions, theoretical papers, and the like. Table 8 shows how each of the article categories was grouped into meta-categories.

Table 8. The Composition of Educational Technology Metacategories.

Qualitative	*Quantitative*	*Mixed methods*	*Other*
Qualitative; critical theory; explanatory descriptive; case studies	Quantitative; experimental/quasi-experimental; quasi-experimental; exploratory descriptive, correlational; causal-comparative; classification; descriptions; experimental research; experimental study; survey research, empirical research; evaluation; correlational; empirical, experimental, or evaluation; quantitative descriptive	Mixed methods; triangulated; mixed	Literature reviews; other; description with no data; theory, position paper, and so forth.; theory; methodology; professional

Note. From *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005*, by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.

What are the Proportions of Quantitative, Qualitative, and Mixed-Methods Research that Educational Technology Researchers Have Tended to Use?

In the previous section, three meta-categories for educational technology research emerged. Figure 8 shows the proportions for each of those categories.

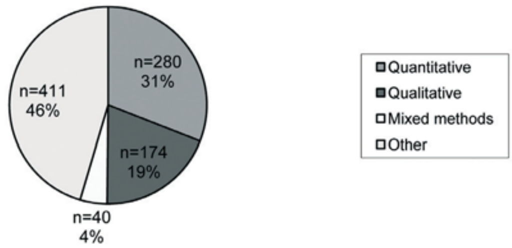


Figure 8. *Proportion of types of articles in educational technology journals. From Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000-2005 (p. 27), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.*

How do those Proportions Differ over Time Periods and Publication Forums?

Figure 9 shows that there is considerable variability across forums in terms of the proportions of *quantitative, qualitative, mixed methods*, or *other* categories of articles. For example, it appears that ETR&D (*Educational Technology Research and Development*) appears to publish papers that do not report on investigations with human participants, while mixed forums tended to publish articles that do report on investigations with human participants.

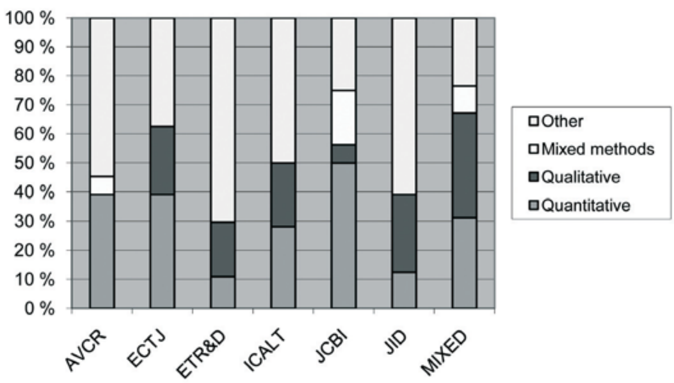


Figure 9. *Proportions of types of educational technology articles by forum. From Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005 (p. 27), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.*

Figure 10 indicates that there was a spike in qualitative educational technology research in the late 80s. However, the spike drops off rapidly and is replaced with quantitative methods as the ‘00s draw near. This finding is also consistent with the declining trend in qualitative research in the field of computer science education (see Randolph 2007a). One contradictory piece of evidence is that Caffarella (1999) found that there was an increasing trend in the proportions of qualitative educational technology dissertations each year from 1977 until 1998.

There is one important caveat about Figure 9 and Figure 10 that should be noted: because of the nature of the methodological reviews included here, time period and forum are confounded (i.e., one cannot separate one from the other). For example, a particular forum may have been sampled more often in a particular time period than other forums, or vice versa. So in reality, it is difficult to say whether it was the time period, forum, or some other confound that was actually associated with the methodology category. However, because the time period trend of a decreasing amount of qualitative research over time (in journals) is consistent with trends in other related fields – like computer science education (see Randolph, 2007a) – the pattern seems to match.

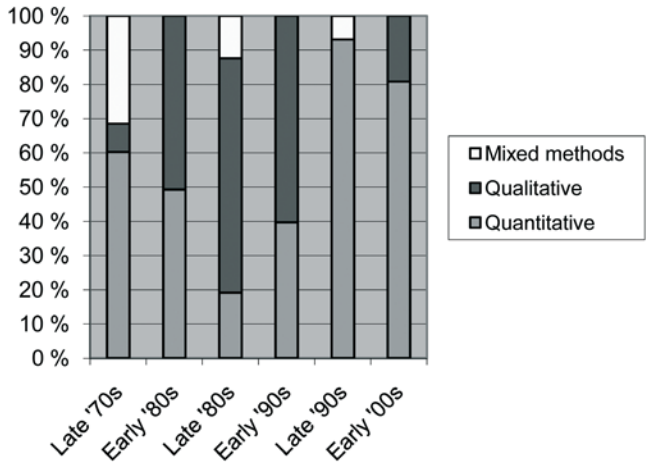


Figure 10. *Proportions of types of educational technology articles by time period. From Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000-2005 (p. 28), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.*

How do those proportions compare to the proportions in the field of education research proper?

In this section, I compare the proportions between the fields of educational technology and education research proper. For the field of educational technology research, I use the data from Randolph's (2007a) synthesis of the methodological reviews of educational technology research that has been discussed in the sections above. For the field of education research proper, I use the data from Gorard and Taylor (2004) who conducted a review of 94 articles – 42 articles from the six issues published in 2001 in the *British Educational Research Journal* (BERJ), 28 articles from the four issues published in 2002 in the *British Journal of Educational Psychology*

(BJEP), and 24 articles from the four issues published in 2002 in *Educational Management and Administration*. Although their sample was small, they provided additional evidence for the generalizability of their results by collecting data from other sources, such as:

> - *interviews with key stakeholders from across the education field, including researchers, practitioner representatives, policy makers and policy implementers;*
> - *a large-scale survey of the current methodological expertise and future training needs of UK education researchers; [and a]*
> - *a detailed analysis and breakdown of the 2001 RAE [Research Assessment Exercise, 2001]. (p. 114)*

Table 9 shows the cross-tabulations of the number of articles, across fields that dealt with investigations with human participants. It shows that education research proper had about a 30% higher proportion of articles dealing with human participants than educational technology research.

Table 9. *Comparison of the Proportion of Human Participants Articles in Educational Technology and Education Proper.*

Field	Human participants		Total	Percentage yes	Adjusted residual
	Yes	No			
Ed. tech	494	411	905	54.6	-5.5
Ed. proper	79	15	94	84.0	5.5
Total	573	426	999		

Note. Ed. tech. = educational technology, Ed. proper = education proper. $X^2(1, N = 999) = 30.21, p < .000$. From *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005* (p. 29), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.

Table 10 shows the cross-tabulations of the number of articles of each methodology type (i.e., quantitative, qualitative, or mixed-methods) across fields. It appears that there were no practically or statistically significant differences between the proportion of quantitative, qualitative, or mixed-methods articles across fields.

Table 10. Comparison of the Type of Methods Used in Educational Technology and Education Research Proper.

Type of article	Field		Total
	Ed. tech	Ed. proper	
Quantitative	280 (56.7%)	43 (54.4%)	323 (56.4%)
Qualitative	174 (35.2%)	32 (40.5%)	206 (36.0%)
Mixed methods	40 (8.1%)	4 (5.1%)	44 (7.7%)
Total	494 (100%)	79 (100%)	573 (100%)

Note. Ed. tech. = educational technology. Ed. proper = education proper. $X^2(2, N = 573) = 1.41, p = .495$. From *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005* (p. 30), by J. J. Randolph, 2007, Unpublished doctoral dissertation, Utah State University. Copyright 2007 by J. J. Randolph.

The conclusion I draw here is that educational technology researchers and education researchers proper tend to do the same proportions of quantitative, qualitative, and mixed-methods research. Where they do differ is that educational technology researchers tend to write more articles that do not deal with human participants. Given that educational technology is a field that deals equally with research and development and that two out of the five main questions in educational technology have to do with theoretical and methodological considerations, it is no surprise then that educational technology research articles less often concern investigations with human participants than their counterparts in education research proper.

In What Proportions do Educational Technology Researchers Choose (a) Research Approaches, (b) Experimental Research Designs, and (d) Measures?

To answer this question I relied primarily on the data presented in Randolph et al. (2005), which is a methodological review of a census of articles published in the 2004 proceedings of ICALT, the International Conference on Advanced Learning Technologies. I verified the generalizability of these findings by comparing the findings of Randolph et al. with findings from other detailed methodological reviews. For example, Randolph et al. (2005) found that 61% (86 of 141) of educational technology investigations did not deal with human participants. This percentage is practically similar to the percentage (i.e., 54%) found in the Randolph (2007a) synthesis of educational technology methodological reviews. Similarly, Randolph (2007a) found nearly the same proportion (66.2%, or 233 of 352) in a related field – computer science education research. Also, the 2004 ICALT proceedings had a 30.5% acceptance rate, which through personal experience is typical for educational technology conferences. I believe that these similarities give at least some preliminary evidence for the generalization from ICALT papers, and papers from the field of computer science education research, to other papers in educational technology.

In terms of the proportions of research approaches used, Table 11 shows the results from Randolph et al.'s (2005) review of ICALT papers. It shows that experimental research, quantitative descriptive research (i.e., survey research), and qualitative research (in descending order) are the most frequently used research methods in educational technology research. The one article in the *Classification* category in the original table was subsumed under the *correlational* category in the table presented here.

Table 11. Research Approaches of ICALT Articles Dealing with Human Participants.

Research approach	*Frequency*	*%*
Experimental/quasi-experimental	21	41
Quantitative descriptive research	13	26
Qualitative	8	16
Correlational	6	12
Causal-comparative	3	6
Total	51	100

Note. From "A Critical Analysis of the Research Methodologies Reported in the Full Papers of the Proceedings of ICALT 2004," by J. J. Randolph, R. Bednarik, P. Silander, J. Gonzalez, N. Mylller, and E. Sutinen, 2005, in *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies* (p. 12), Los Alamitos, CA: IEEE Computer Society. Copyright 2005 IEEE. Reprinted with permission.

Of the 21 articles in Randolph et al. (2005) that used experimental designs, Table 12 shows which and in what proportions the variety of common experimental designs were used. The one-group posttest-only design, the weakest of experimental designs, was used most frequently.

Table 12. Experimental or Quasi-Experimental Designs in ICALT papers.

Design	*Frequency*	*%*
One-group posttest-only	9	43
Pretest-posttest, no controls	4	19
Repeated measures	4	19
Posttest-only, with controls	3	14
Pretest-posttest with controls	1	5
Single-participant	0	0
Total	21	100

Note. From "A Critical Analysis of the Research Methodologies Reported in the Full Papers of the Proceedings of ICALT 2004," by J. J. Randolph, R. Bednarik, P. Silander, J. Gonzalez, N. Mylller, and E. Sutinen, 2005, in *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies* (p. 12), Los Alamitos, CA: IEEE Computer Society. Copyright 2005 by IEEE. Reprinted with permission.

Table 13 shows the measures, and their proportions, for the 21 articles, in Randolph et al. (2005) that used experimental designs. The row totals do not sum to 100% because more than one measure could have been used in each article. As Table 13 shows, student surveys, log files, and teacher- or researcher-made tests were the most common measures used.

Table 13. Measures Used in ICALT Papers.

Measure	*Frequency (of 21)*	*%*
Student questionnaire	19	91
Log files	6	29
Test (teacher/researcher made)	5	23
Interviews with users	5	23
Direct observation	4	19
Exercises	3	14
Teacher survey	2	10
Test (standardized)	2	10
Narrative analysis scheme	2	10
Number of resubmitted exercises	1	5
Time on task (electronic)	1	5
Focus groups	1	5
Pass rate	1	5

Note. From "A Critical Analysis of the Research Methodologies Reported in the Full Papers of the Proceedings of ICALT 2004," by J. J. Randolph, R. Bednarik, P. Silander, J. Gonzalez, N. Mylller, and E. Sutinen, 2005, in *Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies* (p. 13), Los Alamitos, CA: IEEE Computer Society. Copyright 2005 IEEE. Reprinted with permission.

How do Educational Technology Researchers Tend to Report Educational Technology Studies?

Randolph et al. (2005) can shed light on how educational technology researchers report their research. Randolph et al. literally measured the amount of space in each article devoted to (a) reviewing the previous literature, (b) describing the program or intervention, and (c) describing the evaluation (research methods and results) of an intervention. They found that the mean proportions of literature review, program description, and evaluation were 18%, 47%, and 34% respectively; the implication being that educational technology researchers spend a good deal of article space on describing their programs or interventions and spend little article space on reviewing the previous literature or on describing the evaluation of the intervention. These findings parallel the findings from Randolph's (2007a) methodological review of the computer science education literature in which it was found that literature reviews and procedures were grossly underreported.

What Suggestions are Given for Improving Educational Technology Research?

In this section, I summarize and discuss the three major recommendations given in Randolph et al. (2005). The suggestions are summarized in the list below:

- Review and report the previous research.
- Provide sufficient detail in reporting on investigations with human participants.
- Avoid using the one-group posttest-only, attitudes questionnaire.

Review and report the previous research. Randolph et al. (2005) and Randolph (2007a) both converged on the finding that the literature review is given insufficient attention in educational technology research. The theoretical implication is that without a literature review, research questions are unlikely to contribute to information that is important to the scientific community and there can be little hope of developing a cumulative base of knowledge about educational technology and learning design. Also, without literature reviews, developers of educational technology interventions are likely to develop interventions that have already been developed or develop interventions for which there is no need.

Provide sufficient detail in reporting on investigations with human participants. Another finding that can be drawn from Randolph et al. (2005) is that, in general, educational technology articles tend to emphasize a detailed description of the intervention and, as a result, pay insufficient detail to reporting results and procedures of empirical investigations. For example, without an adequate description of procedures, it is difficult or impossible to achieve replication – a cornerstone of science. Also, without an adequate description of settings and participants it is difficult or impossible to establish parameters for generalizing findings.

I suppose that one cause for this insufficient detail is that educational technology, like computing science, is composed of many traditions (e.g., theoretical, engineering, and empirical traditions). In the engineering tradition, the reporting convention is to describe the intervention and its method of construction in great detail (e.g., "I built this thing, which has these specifications"). In the empirical tradition, less attention is given to the intervention description and more attention is given to the results and methods for evaluating the intervention. I hypothesize that the combining of these two traditions has led to hybrid papers that, unfortunately, are insufficient as either engineering or education research papers.

One solution to this problem is to dehybridize educational technology papers. Create one paper that describes the intervention in enough detail that it suffices as an engineering-oriented paper. Create another paper that describes the empirical validation of the intervention. Although this might border on piece-meal publication, it makes sense in a field characterized by multidisciplinarity and whose main reporting format is the five-page conference paper.

Avoid using the one-group posttest-only, attitudes questionnaire. The one-group posttest-design should be avoided because it is subject to almost all threats to validity. For example, suppose that an educational technology researcher implemented an intervention in a course and gave out a questionnaire to students about whether they felt that their learning had increased during the time that they were in the course. First, there are many things that probably had an effect on reported learning other than the educational technology intervention. Perhaps, the teacher provided effective instruction, the textbook was sufficient, the students studied outside of the course, among many other possible causal explanations.

While measuring attitudes (in this case, self-reports of learning) is an easy way to collect learning data, self-reports of learning are historically unreliable (Almstrum, Hazzan, Guzdzial, & Petre, 2005). Also, they suffer from what Shadish, Cook, and Campbell (2002) call *mono-operation bias* and *mono-method bias*.

Questionnaires are ideal for collecting large amounts of data so that generalizations can be made from a sample to population. However, in the cases when the research priority is not to make generalizations, but rather to provide thick description of a phenomenon, measures other than questionnaires might be more appropriate. For example, in the often-seen case when an educational technology researcher implements a new intervention in his or her own classroom and the research goal is to collect information that can be used to improve the intervention, it might be better to conduct in-depth interviews with students or observe log files to get information that is useful for development. Knowing for example that "85% of students reported that they liked the intervention a little or a lot," provides little or no information about how to improve the intervention.

Summary

In this section I summarize the results of some of the major methodological reviews of the educational technology research to answer the research questions listed below:

1. What are the meta-categories that can be used to subsume the research categories in other methodological reviews of educational technology research?
2. What are the proportions of quantitative, qualitative, and mixed methods research that educational technology researchers have tended to use?
3. How do those proportions differ over time periods and publication forums?

4. How do those proportions compare to the proportions in the field of education research proper?

5. In what proportions do educational technology researchers choose (a) research approaches, (b) experimental research designs, (c) outcomes, and (d) measures?

6. How do educational technology researchers tend to report educational technology studies?

7. What suggestions are given for improving educational technology research?

A brief summary to each of those questions is given below:

1. One set of meta-categories that can be used to subsume the research categories in other methodological reviews are *quantitative, qualitative, mixed-methods* and *other*. The *other* category consists of articles that do not deal with investigations with human participants, such as literature reviews, theoretical papers, or program descriptions.

2. Articles that do not deal with human participants make up 46% of the articles published in educational technology research. Of the articles that do report on investigations involving human participants, the percentages of quantitative, qualitative, and mixed methods research are 56.7%, 35.2%, and 8.1%.

3. There is considerable variability in what type of methods the major journals in educational technology tend to publish. Concerning research trends, it appears that there was a spike of qualitative research in the late 1980s and a spike in quantitative research in the late 1990s and early 2000s. Note, however, that year and publication forum are confounded in these results.

4. While educational technology research tends to have more articles that do not deal with human participants than in education research proper, the proportions of quantitative, qualitative, and mixed-methods research is nearly the same.

5. Educational technology researchers tend to use (in descending frequency) experimental, quantitative descriptive (i.e., survey research), and qualitative research the most. When they do experimental research, they tend to use the one-group posttest-only design. In decreasing order of frequency, they tend to use questionnaires, log files, and teacher- or researcher-made tests as measures and the outcomes of interest are typically student attitudes, academic achievement, and attendance.

6. Educational technology researchers tend to describe their interventions in great detail. However, they also tend to provide inadequate literature reviews

and inadequate detail in their reports of research done with human participants.

7. Recommendations for improving educational technology research include reviewing and reporting the previous research, providing sufficient procedural and contextual research details, and avoiding the one-group posttest-only attitudes questionnaire.

Questions to Consider

1. Coming back to the table about the proportions of quantitative, qualitative and mixed-methods research in educational technology (Table 10), the proportions were 56.7%, 35.2%, and 8.1%, respectively. What conclusions about the disciplinary identity of the field can you draw, if any, from those findings?

2. In the field of computer science education, Randolph (2007a) found that nearly 40% of articles dealing with human participants only presented anecdotal evidence for their claims. Do you think that the proportion of articles that present anecdotal-only evidence could also be that high in the field of educational technology?

3. Why do you think that educational technology research has a higher proportion of articles that do not involve human participants than in education research proper?

4. What is your opinion about the posttest-only attitudes questionnaire? What is its place in educational technology research?

Chapter 4

Data Collection in Educational Technology Research

In the previous chapter, I presented information about the types of data that educational technology researchers tended to collect and what measurement instruments they tended to use. In this chapter, I will discuss in more detail the major methods of data collection and also point out how technologies can be used in the collection of research data.

To help frame this discussion, I will use the framework for valuing data collection methods mentioned in Frechtling (2002). In Frechtling's framework, two issues are important in analyzing data collection methods: theoretical issues and practical issues.

The theoretical issues involve (a) the value of the data generated, (b) the scientific rigor of the data generated, and (c) the philosophical issues underlying the data collection method. Concerning value, different types of data can have more or less value to a researcher depending on the research situation. For example, data that indicate that "80% of students reported that they liked an intervention 'a little' or 'a lot'," do little for a researcher trying to determine how to improve an intervention or understand why an intervention works. On the other hand, that same data might indeed prove useful for a researcher trying to investigate whether students, in general, are satisfied with an intervention. Put in another way, sometimes data with depth are preferred to data with breadth, or vice versa.

Concerning rigor, data collection methods vary in the degree of rigor that is required. For example, anecdotal observations have little rigor while single-participant observations, which involve carefully operationalizing variables and using multiple observers, have much rigor.

In terms of philosophical paradigms, some types of data collection lend themselves to some paradigms more than others. For example, if a researcher bases his or her research on the idea that there is not an objective reality and that reality is socially constructed, then it would not seem fitting to use an instrument that purports to measure objective states of reality.

The practical issues in data collection involve (a) the credibility of the findings, (b) staff skills, (c) costs, and (d) time constraints. Concerning credibility, different types of data have different credibility for different audiences. Also, the variety of data collection methods differ in the amount of staff skills that are needed to do the data collection, the costs involved with the data collection, and the time needed to collect the data. For example, questionnaires are usually easily and rapidly administered and cost relatively little. On the other hand, conducting ethnographic research takes much preparation on the part of the researcher and a commitment to spending long and probably costly amounts of time in the field.

Since it was found that questionnaires were the most frequently used instrument in educational technology research, it is fitting that questionnaire data collection methods are dealt with first. In order of frequency of use, I then go onto discuss the four other most frequently used data collection methods in educational technology: log files, tests, interviews, and direct observation.

Questionnaires

Questionnaires are ideal for quantitative descriptive research when data need to be collected from many people and when the questions to be asked are clearly defined. They are ideal for research that is meant to have breadth, rather than depth. (I use the term *questionnaire* synonymously with *survey*.) For example, if the goal is to find out the level of student satisfaction with an intervention, then a questionnaire might be appropriate. However, if the goal is thick, rich description, then other data collection methods, like interviews, might be more appropriate.

According to Frechtling (2002) the advantages of using questionnaires are that they are good for collecting basic descriptive data, they are inexpensive, and the quantitative data generated from them can be easily imported or transcribed into statistical software. On the other hand, questionnaires are most often based on self-reports, which are often biased, and their data usually lack depth and do not take context into account (Frechtling, 2002). Also, questionnaires are vulnerable to response bias – bias that occurs when the individuals who responded to a survey differ in important ways from those who did not respond. Typical problems are that people who drop out of an intervention tend not to get a postintervention satisfaction survey and that people who fill out surveys tend to be different than people who do not fill out surveys (Gall, Borg, & Gall, 1996). It is a much different claim to say that "80% of all of the people who tried the intervention reported that they liked it 'a little' or 'a lot'" than "80% of the people who tried the intervention and liked it enough to finish it and liked the researcher enough (who also designed the intervention) to fill out a questionnaire, reported that they liked the intervention 'a little' or 'a lot.'" One way to deal with response bias is to assess the ways that the people who did not respond to the survey differed from the people who did respond to the survey (Gall, Borg, & Gall, 1996). Dillman (1978) provides a wealth of information about how to improve the response rate of surveys, including offering rewards, sending follow-ups, and so on.

Questionnaires can have open and closed questions and can be administered in many ways. In closed questions, respondents are asked to select from a set of fixed responses. Many questionnaires use what is called a *Likert scale* in which respondents select one response in a ranked series of responses. For example, respondents might be asked to select the degree to which they agree with a certain statement. The levels of agreement might be: *strongly disagree, disagree, neither agree nor disagree, agree*, or *strongly agree*. In open questions, respondents are allowed to give a free-form answer. For example, a respondent might be expected to write in what they liked or disliked about an intervention. The term *questionnaire* connotes that the questions be administered in writing; however, it is possible to use questionnaires in face-to-face interviews, in telephone interviews, or through computers. In fact, much research lately has gone into the differences in the quality of data generated from paper and pencil and online questionnaires. One consistent difference is that respondents tend to disclose more information and give less socially desirable answers in electronic surveys than in paper-based mail surveys (Kiesler & Sproull, 1986). See De Leeuw (1993) or Dilman (1978) for a theoretical model and empirical research on the effects of different survey media on data quantity and quality.

Also, there is a growing field of technology related to using computing technology for questionnaires and interviews. For example, VIRRE, which is portrayed in Figure 11, is a computer-assisted self-interview tool, designed for administering questionnaires or interviews to youths. The survey question appears on the screen in VIRRE's abdomen and the students can respond by pushing the appropriate button on VIRRE's screen or by recording their response using a webcam. Randolph, Virnes, Jormanainen, and Eronen (2006) found that VIRRE leads to more interview data than when data are collected using a pencil-and-paper interview format. A good review of computer-assisted interview tools can be found in De Leeuw and Nichols (1996), Saris (1989), or Saris (1991).

Figure 11. *Virre: A computer-assisted self-interview tool. From "The Effects of a Computer-Assisted Interview Tool on Data Quality," by J. J. Randolph, M. Virnes, I. Jormanainen, and P. J. Eronen, 2006, Educational Technology & Society, 9(3), p. 199. Copyright 2006 by the International Forum of Educational Technology & Society (IFETS). Reprinted with permission.*

Questionnaires, whether computer-assisted or not, can vary substantially in their rigor. When accuracy and rigor are important, a researcher might need to go through the long process of doing a reliability or validity study of the questionnaire (Spector, 1992). In other cases, it might be enough just to do some pilot testing and then administer the study. At any rate, designing a questionnaire, especially a reliable one, is more difficult and time consuming than it might seem. Even a simple questionnaire will need probably a few iterations of pilot testing and development. A reliability study and item analysis needs about 150 respondents per iteration (Spector, 1992). Therefore, before attempting to make a survey oneself, I suggest trying to find a pre-existing survey. Many educational technology surveys already have reliability and validity information provided for them. For example, the Texas Center for Educational Technology's website (n.d) provides many useful surveys regarding attitudes towards computers. Other places to find pre-existing surveys are in published dissertations or theses, in academic databases (such as ERIC), or, of course, on the Internet. The Educational Testing Service (n.d., n. p.) also maintains an online "library of more than 25,000 tests and other measurement devices that makes information on standardized tests and research instruments available to researchers, graduate students, and teachers. Collected from the early 1900s to the present, the Test Collection at ETS is the largest such compilation in the world."

For those unlucky enough not to find a pre-existing survey to suit the purpose or those brave enough to create one on their own, there are many good resources. Unfortunately, a detailed synthesis of survey construction and testing is beyond the scope of this text; however, there are many great resources – the most notable of which is Arlene Fink's 1,434-page, 10 volume, *Survey Kit* (2002).

Log Files

Educational technology researchers have a unique advantage over other types of researchers: educational technology interventions often have automated ways of recording user-generated data. That user-generated data is a boon for educational technology researchers because it is recorded unobtrusively and in a naturalistic setting. Also, because the process is automated, loads of data can be collected at almost no expense.

Examples of log files could be records of when and how long students used a particular intervention, what features they used, how often they used those features, or how often they used one feature in conjunction with another feature. Other examples of log files could be texts from chat rooms or bulletin boards (Rieger & Sturgill, 1999). Log files could even be used to create maps of social networks (Durland & Fredericks, 2006). Log files could also consist of the results of automatically graded student exercises (as in Laakso, Salakoski, Korhonen, & Malmi, 2004). In essence, any aspect of users' interaction with the learning environment or with each other that can be recorded can become a rich source of data. (I use a broad definition of log files that includes things like screen recorders too.) Rieger and Sturgill (1999) provide a framework and a review of the tools for observing users and gathering feedback about electronic environments. Some of the tools they mention are log files, screen recorders, persistent cookies, agents, automated services (e.g., Google analytics), and user artifacts. Eye-tracking is yet another emerging technology for logging the behaviors of participants as they interact with learning environments

(Bednarik & Randolph, in press). Gay and Bennington (1999), which is the edited volume that contains Rieger and Sturgill (1999), would be of interest to educational technology researchers interested in unobtrusive ways of learning about their interventions.

Some of the benefits of using log files are that data collection is unobtrusive and that the data collected are naturalistic. The participants are free to work as they normally do while data are being collected; there is no need to set up artificial situations or environments. There is no researcher peering over a participant's shoulder. That is not to say that the participants do not need to know that their activities are being recorded. To do ethical research, the participants must give their permission for their data to be recorded and to be used for research purposes. One disadvantage of using log files is that they often create overwhelming amounts of data.

Tests

Researcher-made or teacher-made tests, in contrast to standardized tests made by a testing company, were the third most frequently used measure in educational technology research. Tests are useful when a researcher wants to document the state of knowledge at one time or wants to measure how knowledge changes over time or under a variety of conditions. Tests have the benefit that they can provide "objective" information on what a participant knows or can do, they are relatively easy to score (unlike performance measurements), and are generally accepted by the public as an indicator of learning (Frechtling, 2002). On the other hand, there are also many disadvantages to testing. They might measure a construct other than the one they are intended to measure (Shadish, Cook, & Campbell, 2002), they might measure knowledge unreliably, they might be biased towards certain populations, or they can be corrupted by teaching or coaching, among other disadvantages.

Some benefits of teacher-made tests are that they are easy to construct, inexpensive to administer, and not outside of the natural contexts of what students normally do in the classroom. However, teacher-made tests lack some of the desirable features that carefully-made (and usually expensive and unnatural) standardized tests do, such as empirical information about a test's reliability and validity. As in methods choice in general, there is always a trade-off in choosing what types of measures to use.

Like the creation of surveys, the creation of reliable and valid tests of knowledge is very difficult; in fact, there is a whole field of research, *psychometrics,* dedicated to it. Therefore, it is often best to identify a pre-existing test if it is possible – again, check previous dissertations, academic databases, the Internet, and test collections. In research that requires much rigor, information about the reliability of a test is usually needed. A classic text on testing and test creation is Thorndike (2004).

For those displeased with traditional testing (see Wiggins, 1993), another option is to concentrate on using performance measurements rather than traditional pencil and paper tests. Performance-based assessments like portfolios, student demonstrations, student products, and problem-solving activities are more naturalistic, measure different types of knowledge than traditional tests do, and often require the integration of a broad range of knowledge (Glatthorn, Bragaw, Dawkins,

& Parker, 1998). Performance-based assessments have the disadvantage that they have more subjectivity in their scoring than traditional tests and that their scoring is not as straightforward, and, therefore, not as amenable to quantitative analysis, as traditional tests.

Technology has done much to advance the field of testing. For example, nowadays computers can be used to rate academic essays to good effect (Rudner & Gagne, 2001), to provide individualized feedback on writing improvement (Chen, 1997), and to adapt to the level of skill of a test taker (computer-adaptive testing) (Wainer et al., 2000). See Alderson (2000) for a review of the future of technology in testing.

Interviews

While surveys are ideal for collecting shallow knowledge over a broad area, interviews are ideal for collecting deep knowledge over a limited area. In addition to generating rich data, they allow for face-to-face contact, enable a researcher to follow up immediately on unclear or ambiguous answers, gain access to information that a respondent would not reveal on paper, and give the researcher flexibility in administering the interview according to the needs of individual respondents or situations. Patton, as cited in Frechtling (2002), gives a list of example evaluation questions that are appropriate to answer using interviews:

- *What does the program look and feel like to the participants? To other stakeholders?*
- *What do stakeholders know about the project?*
- *What thoughts do stakeholders knowledgeable about the program have concerning program operations, processes, and outcomes?*
- *What are participants' and stakeholders' expectations?*
- *What features of the project are most salient to the participants?*
- *What changes do participants perceive in themselves as a result of their involvement in the project? (p. 51)*

While interviews have many advantages, they are time-consuming to conduct and the interviewers need considerable training. Also, a great deal of time usually needs to be spent transcribing interviews. Also, because of the flexibility of open interviews, there is a possibility for inconsistencies between different interviewers. Because of the face-to-face contact between interviewers and respondents, there is the possibility for interviewers to lead respondents. Finally, the volume of information that is generated can be very large and overwhelming.

Interviews can vary to the degree that they are planned a priori. In some cases, the interview protocol can be very strict and respondents are only allowed to choose from a certain set of responses (i.e., a verbal questionnaire). In other cases the interview might be less structured and the interviewer has much freedom to explore and follow up on a participant's responses. In some types of interviews, the interviewer records the responses on the spot. In other types of interviews, the interviewer painstakingly transcribes audio or video recordings. In discourse analysis (see Potter & Wetherell, 1987), in addition to words, the researcher transcribes pauses, elongations and emphases of syllables, overlaps between speakers, and other features of spoken language. There are many good guides for collecting interview data. Some of those I recommend are Rubin and Rubin (2004) and Kvale (1996).

Although not as often as interviews, focus groups are sometimes used in educational technology research. Essentially, a focus group is an interview with 8-12 people who share common characteristics. Originally used in market research, the focus group technique capitalizes on the group dynamic to generate insights that would not have been generated if the participants had been interviewed individually. According to Frechtling (2002), focus groups are useful for "identifying and defining problems in project implementation; pretesting topics or ideas; identifying project strengths, weaknesses, and recommendations; assisting with interpretation of qualitative findings; obtaining perceptions of project outcomes and impacts; [and] generating new ideas" (p. 53). Frechtling (2002) provides many helpful recommendations, which are too numerous to report here, for choosing between focus groups and interviews. A helpful resource for conducting focus groups is Stewart, Shamdasani, and Rook (2006).

There are many technologies available to aid interviewing. Two of the major types of technologies are computer-assisted telephone interviewing [CATI] and computer-assisted personal interviewing [CAPI] (de Leeuw & Nichols, 1996; Saris, 1989, 1991). The latest generation of interview tools totally automates the personal interviewing process (see, e.g., Randolph, Virnes, Jormanainen, & Eronen, 2006).

Direct Observation

Direct observation, where a researcher directly observes the behaviors of participants, is another form of data collection used often in educational technology research. It comes in many forms. In one form, direct observation can be as simple as peering over a participant's shoulder and noting how a participant interacts with a learning tool over time. In another form (as is the tradition in single-participant research), the researcher carefully operationalizes a behavior or set of behaviors and counts the number of times that the behavior occurs over a given period. For example, an educational technology researcher might observe how often students interact with each other when participating in a technologically oriented intervention (Sutinen, Virmajoki-Tyrväinen, & Virnes, 2005). To establish estimates of reliability in this type of research, additional researchers might also make observations to see to what degree two or more observers can agree on whether a specific behavior occurred (see Kazdin, 1982).

Another form of direct observation, in the ethnographic tradition, is to observe and participate in a culture-sharing group (see, e.g., Cronjé, 2006). An increasingly seen form of direct observation, albeit of a different character than the others, is verbal protocol methodology, "a data collection technique that asks the user-consumer to talk aloud while interacting with the technology, thus revealing the cognitive processes of the user" (Mathison, Meyer, & Vargas, 1999, p. 74). Although this form of observation can help explain the cognitive processes of users, its main shortcoming is that it is obtrusive. Having to explain what one is thinking while performing an activity, can influence how one does the activity

Observational techniques are advantageous because they allow the researcher to get first-hand experience of the behavior of participants, individually and as a group. Also, they allow the researcher to come to understand the contexts important to the research situation. Frechtling (2002) argues that observations "may be especially important where it is not the event that is of interest, but rather how that event may fit into, or be affected by, a sequence of events" (p. 54). On the other hand, direct observations are time-consuming and often require much training on the part of the observer. Also, the act of observation often changes that which is being observed.

The benefits of creating audio/video records of observations are many. Their permanency helps bolster the validity of findings. If there is a dispute about a claim, the recordings can be used as solid evidence. Second, recording an observation allows a researcher to replay findings and, thus, enhance the accuracy of findings or gain insights that might have been missed had the recording not been reviewed. However, recording behavior has a strong tendency to influence behavior – at least, until the participants become accustomed to being recorded.

Mixing Data Collection Methods

There are many good reasons for using a variety of data collection methods when conducting educational technology research, the most of important of which is that combining methods, what Frechtling (2002) calls *triangulation,* increases the validity of findings. As I have pointed out, all of the data collection methods mentioned above have their strengths and weaknesses. By combining techniques, one can cancel out the weakness of one technique by pairing it with a complementary technique. For example, the observation of behavior might be paired with interviews that expose the intentions behind those behaviors to gain a holistic picture of the cognitive and behavioral processes involved. Another reason is that different stages of research call for different methods of investigation. For example, Frechtling (2002) argues that one might begin an investigation with a qualitative focus group approach to identify issues for further investigation, then conduct a survey to see to what degree that issue affects the population of interest, and conclude by conducting another qualitative inquiry to gain deeper insights into the issue. There are several good resources on how to mix methods to a researcher's advantage, the most prominent of which are Tashakkori and Teddlie's (2002) *Handbook of Mixed Methods in Social & Behavioral Research,* Cresswell and Plano Clark's (2006) *Designing and Conducting Mixed Methods Research,* and Johnson and Onwuegbuzie's (2004) seminal article – *Mixed methods research: A research paradigm whose time has come.*

Questions to Consider

1. What data collection methods have you used in the past? What advice would you give to a colleague who intends to use one of those data collection methods for the first time?
2. Of the data collection methods mentioned above, which has the most personal appeal to you?
3. I mentioned the advantages of using multiple sources of data collection. Can you anticipate some of the disadvantages? In particular, what happens when data collected in two different ways conflict with one another?
4. I highlighted some of the technologies that are currently being used to aid in data collection. Can you think of other ways that technologies can be used, or are being used, in educational technology research?

Chapter 5

The Analysis of Educational Technology Research Data

As discussed in the previous section, the data generated from educational technology research can take numerous forms: log files, eye-tracking fixations, test results, interview transcripts, direct observations, talk-aloud transcripts, survey data, and many others. Accordingly, there is a large variety of approaches to analyzing the numerous forms of educational technology data.

I have found it helpful to group the approaches of data analysis into four categories:

1. The quantitative analysis of quantitative data,
2. The quantitative analysis of qualitative data,
3. The qualitative analysis of qualitative data, and
4. The qualitative analysis of quantitative data.

The goal of this section on data analysis is simply to acquaint the reader with the different methods of analysis available and to refer readers to key resources for each method.

The Quantitative Analysis of Quantitative Data

The quantitative analysis of quantitative data is complex and has a subfield of mathematics devoted to it – statistics. The field of statistics can prove to be very useful for making sense of complex data, gaining insights into phenomena, making accurate predictions, making inferences from a population to a sample or from a sample to a population, making classifications and associations, or even revealing and testing theory, however, how to do statistics is a discussion far outside the scope of this text. While not every educational technology researcher may be inclined to do quantitative analyses, I believe that a well-rounded researcher at least needs to be able to interpret the results of other researcher's quantitative reports. There are several good resources that meet this end (e.g., Huck, 2007; Thorndike & Dinnel, 2000).

The Quantitative Analysis of Qualitative Data

While at first, the term *quantitative analysis of qualitative data* might seem awkward, it is an increasingly seen method of analysis in educational technology research. One of its most popular manifestations is quantitative content analysis. Neuendorf's (2002) definition of (quantitative) content analysis is "the systematic, objective, quantitative analysis of message characteristics" (p. 1). Quantitative content analysis originated in media studies, but has been increasingly adopted by educational technology researchers. While content analysis is often done with texts, it can be used actually with any sort of media – websites, movies, pictures, songs, and so forth.

In contrast to the qualitative analysis of qualitative data, according to Neuendorf (2002) quantitative content analysis (a) adheres more closely to the scientific method and is closer to the positivistic paradigm, (b) emphasizes objectivity-intersubjectivity (intersubjectivity is the standard that relates to how a group of people can agree that something is true), (c) uses an a priori design, (d) stresses reliability and validity, (e) creates generalizable results, (f) is replicable, and (g) is amenable to hypothesis testing.

Researchers use *quantitative content analysis* to quantify the degree to which a certain category or event occurs in a sample or population. This quantified information can then be used for comparative, correlational, or even experimental research. In contrast, researchers use *qualitative content analysis* to specify what categories or events occur, why or how they occur, in what contexts they occur, or the meanings of their occurrences. In many ways, quantitative and qualitative content analysis are complementary. For example, a qualitative study could be used to identify the variables underlying a phenomenon. A quantitative content analysis study, using the categories identified in the qualitative study, could then be used, to ascertain to what degree those categories are present, how they quantitatively relate to each other, how they change over time, or how they relate to other variables.

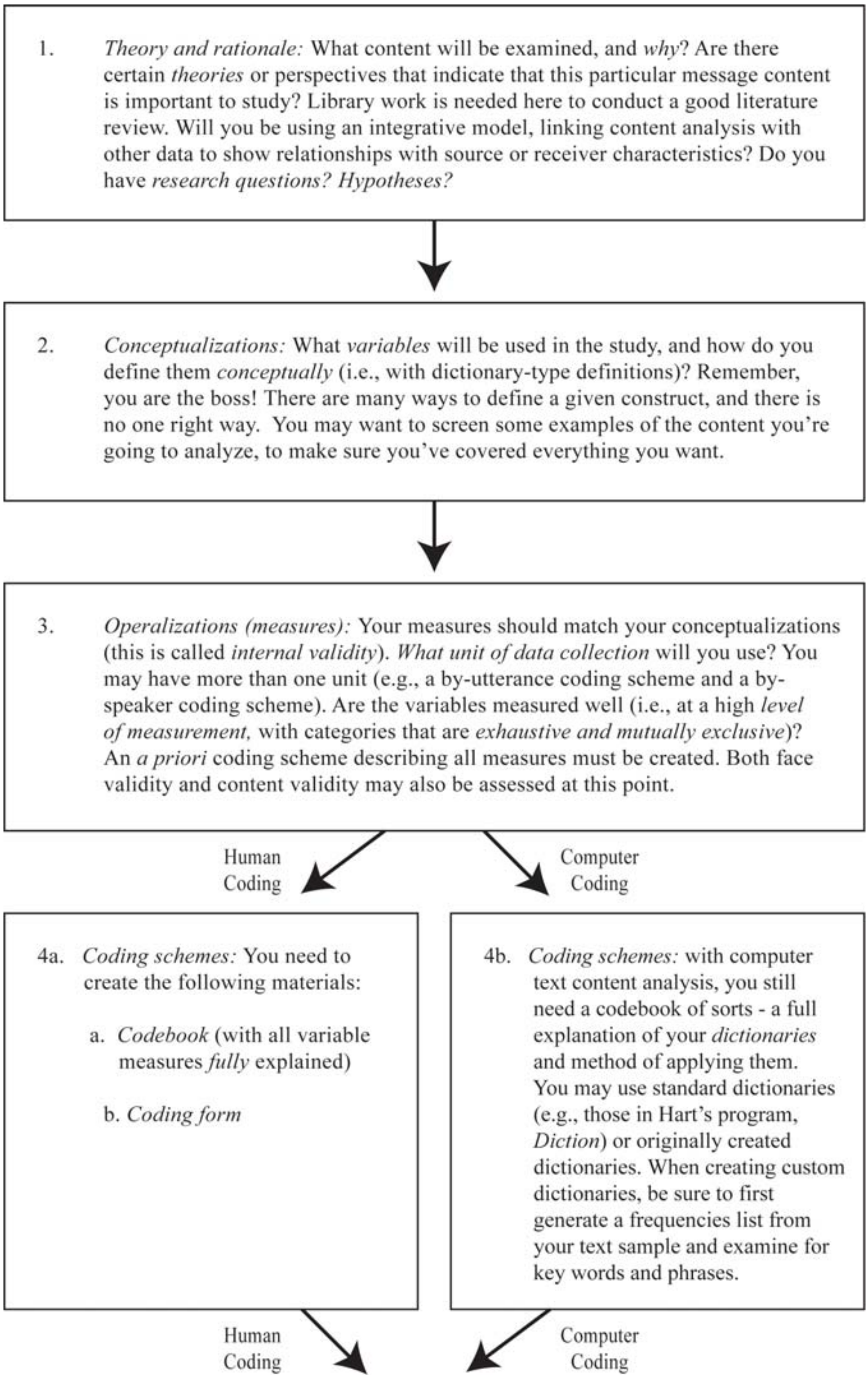
Neuendorf (2002) suggests that there are nine steps in conducting a human-coded quantitative content analysis, which parallel the steps, more or less, of many other types of research. Figure 12 illustrates the nine steps in Neuendorf's brand of content analysis. The first step is to establish the theory and rationale for a study. The second step, conceptualization, involves identifying variables and defining them conceptually. The third step is to operationalize those variables identified in the second step. The goal is to find measures that match the variables' conceptualizations. One can use an a priori coding scheme (i.e., a coding scheme developed by someone else) or one can use an emergent coding scheme (i.e., a coding scheme based on categories that emerge along the way). This step also includes choosing a unit of analysis (e.g., a phrase, a sentence, or a speaking turn). With human coding, the fourth step is to create a coding book and coding form. In the coding book, one explains the variables in great detail and the procedures for coding the data. There should be enough detail and so little ambiguity that two independent coders could consistently apply the same codes to the same units of analysis. The coding form is a form on which the coder inputs the data. (With computer coding, the fourth step is to find an existing dictionary or to create a custom dictionary.) The fifth step is to choose a sample. If the content to be examined is manageable, one might choose to do a census and code all of the content. If there is a large amount of content to code, one might choose a random sample (or some other type of sample) to code. The sixth step is to train coders in the coding book and coding procedure and, then, to pilot test and refine the coding procedure until a sufficient level of reliability is achieved. The seventh step is to do the actual coding of the content and the eighth step is to calculate final reliabilities for the coding. The ninth and final step is to tabulate and report the data. One might also do statistical analyses to identify relationships between content analysis variables, to identify changes in content analysis variables over time, or to identify relationships between content analysis variables and other, external variables. Neuendorf (2002) and Krippendorf (1980) are excellent resources for those interested in quantitative content analysis. Osorio (1998) discusses the use of content analysis for analyzing transcripts in online courses.

The Qualitative Analysis of Qualitative Data

There are many popular methods for the qualitative analysis of qualitative data (e.g., Merriam, 2001; Miles & Huberman, 1994; Madison, 2005; Wolcott, 1994) all of which are basically variations of what Creswell (2007) calls the *data analysis spiral*. Creswell calls it a *spiral* because "the process of data collection, data analysis, and report writing are not distinct steps in the process – they are interrelated and often go on simultaneously in a research project" (p. 150.) Figure 13 is an illustration of Creswell's data analysis spiral.

The first step in the spiral is data management. In this step, researchers segment, document, and organize their data. They segment their dataset by breaking it down into its units of analysis (e.g., sentences, paragraphs, a speaking turns, etc.). They can also document and organize their data, using the same logic that archaeologists use to document their artefacts, to create what Yin (2003) calls a case study database.

After setting up a system to manage the data, qualitative researchers do some form of reading and memoing. The researcher begins by sifting through the documents, reading and reflecting on them. After coming to understand the content, the researcher might begin to write down memos in the margins about concepts that occur to the researcher (Merriam, 2001) or highlight meaningful statements in the texts (Moustakas, 1994).


1. Theory and rationale: What content will be examined, and why? Are there certain theories or perspectives that indicate that this particular message content is important to study? Library work is needed here to conduct a good literature review. Will you be using an integrative model, linking content analysis with other data to show relationships with source or receiver characteristics? Do you have research questions? Hypotheses?
2. Conceptualizations: What variables will be used in the study, and how do you define them conceptually (i.e., with dictionary-type definitions)? Remember, you are the boss! There are many ways to define a given construct, and there is no one right way. You may want to screen some examples of the content you're going to analyze, to make sure you've covered everything you want.
3. Operalizations (measures): Your measures should match your conceptualizations (this is called internal validity). What unit of data collection will you use? You may have more than one unit (e.g., a by-utterance coding scheme and a by-speaker coding scheme). Are the variables measured well (i.e., at a high level of measurement, with categories that are exhaustive and mutually exclusive)? An a priori coding scheme describing all measures must be created. Both face validity and content validity may also be assessed at this point.
Human Coding
Computer Coding
4a. Coding schemes: You need to create the following materials:
a. Codebook (with all variable measures fully explained)
b. Coding form
4b. Coding schemes: with computer text content analysis, you still need a codebook of sorts - a full explanation of your dictionaries and method of applying them. You may use standard dictionaries (e.g., those in Hart's program, Diction) or originally created dictionaries. When creating custom dictionaries, be sure to first generate a frequencies list from your text sample and examine for key words and phrases.
Human Coding
Computer Coding

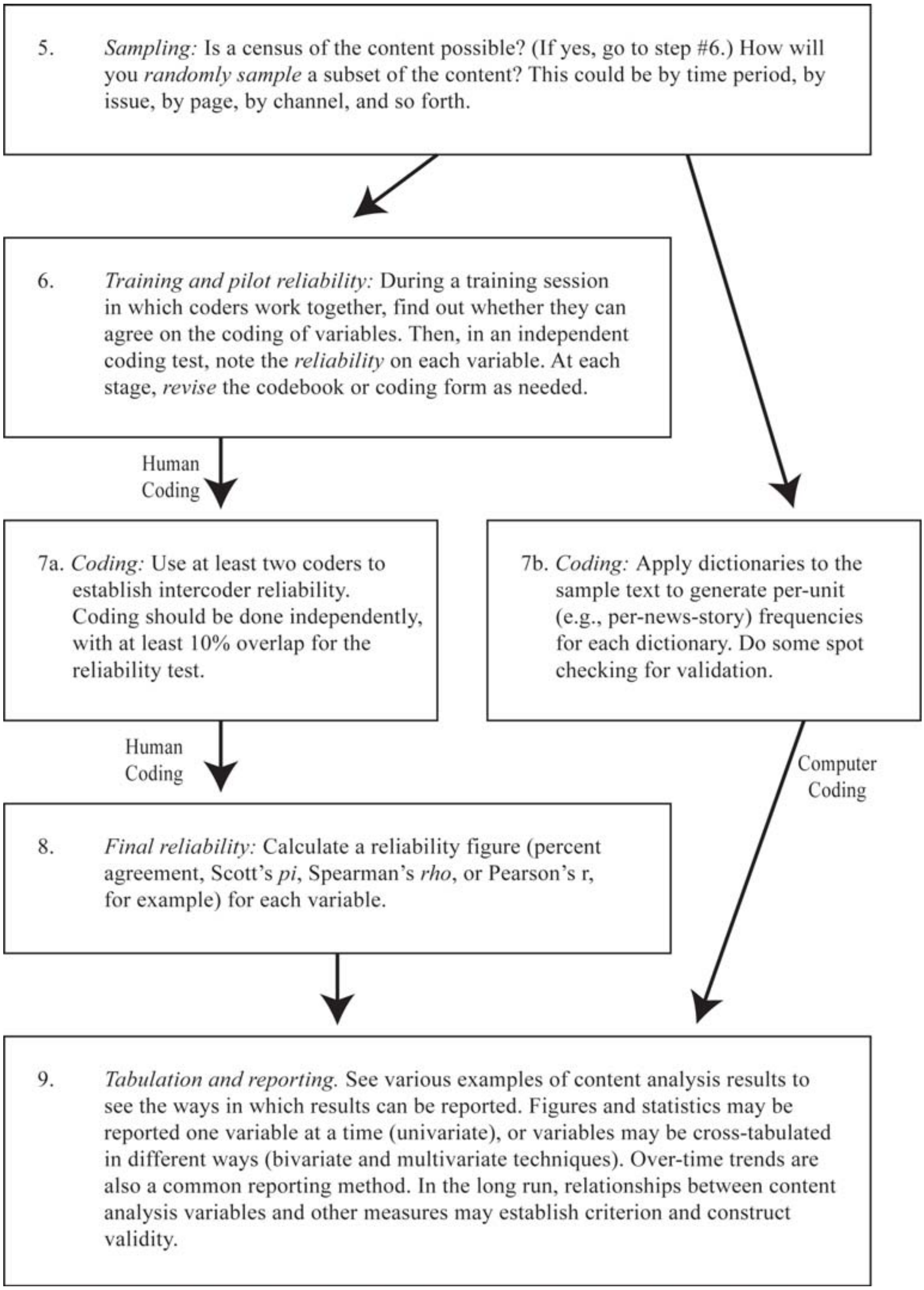


Figure 12. *A flowchart for the typical process of content analysis. From The Content Analysis Guidebook (pp. 50-51), by K. A. Neuendorf, 2002, Thousand Oaks, CA: Sage. Copyright 2002 by K. A. Neuendorf. Reprinted with permission.*

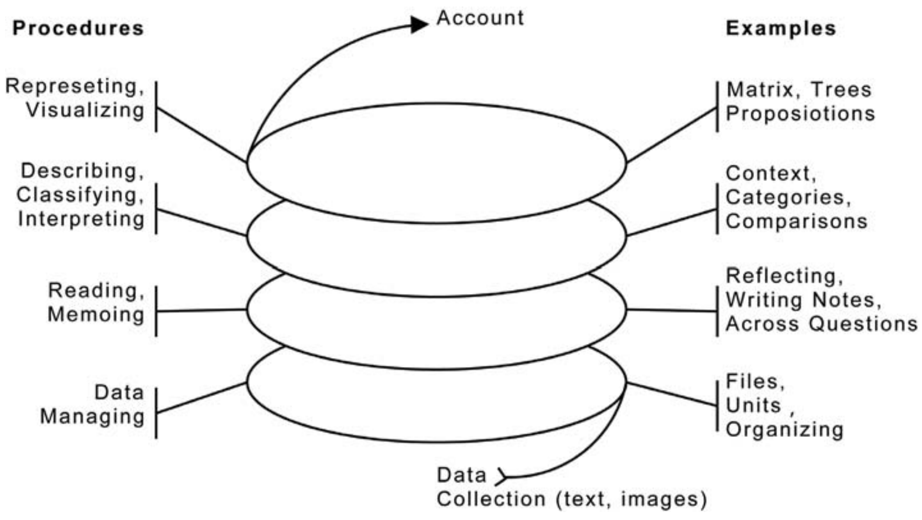


Figure 13. *The data analysis spiral. From Qualitative inquiry & research design: Choosing among five approaches (2nd ed.) (p. 151), by J. W. Creswell, 2007, Thousand Oaks, CA: Sage. Copyright 2007 by Sage Publications, Inc. Reprinted with permission.*

The next step in Cresswell's (2007) spiral is to do description, classification, and interpretation. In this step, the researchers might begin by providing detailed description. According to Cresswell, "this detail is provided *in situ*, that is, within the context of the setting of the person, place, or event. Description becomes a good place to start in a qualitative study (after reading and managing data), and it plays a central role in ethnographic and case studies" (p. 151). Then the researcher attempts to develop an initial set of categories for the data. Some researchers begin with a short list of general categories, and then in later iterations break the codes into increasingly specific categories. Others use the opposite approach – beginning with many specific categories and synthesizing them into general categories.

After the categories and subcategories for the data are established, the next step is to interpret or make sense of the data. Researchers might for example posit theories about a phenomenon or identify relationships between persons, setting, and events.

The next step is to find ways to represent the data. This might include creating matrices of data, hierarchical family trees, causal illustrations, social network maps, or figures. At this point the researcher might present preliminary findings to stakeholders to get their interpretations of the data and to increase the validity of the data through member checking.

The final step in the spiral is to write up one or more versions of the research report. I will discuss this step in more detail in the next chapter.

The prominent books on the analysis of qualitative data include Miles and Huberman (1994), Merriam (2001), and Potter and Wetherell (1987). Also, there are several qualitative computer programs that can simplify the process of qualitative analysis, if a researcher finds it worthwhile to invest the time to learn how to use them. A thorough, but dated, review of 24 of these programs can be found in Weitzman and Miles (1995) and a more recent review in Creswell (2007).

The Qualitative Analysis of Quantitative Data

Although the notion of *qualitative analysis of quantitative data* at first might seem awkward, in fact, it is an everyday occurrence in quantitative research. (It might be more accurate to refer to this method of analysis as *the visual analysis of quantitative data*.) For example, creating a graph of quantitative data is an example of qualitatively analyzing quantitative data.

Qualitative analysis of quantitative data is the main method of analysis of single-participant research. A proverb in single participant research is that, "if you cannot see the effect, it is probably not there." For example, coming back to Figure 7, through a visual analysis of the data it becomes clear that the scores during the intervention increased.

Mixed-Methods of Analysis

In reality, there is not a clear distinction between these methods of analysis. For example, some qualitative data analysts, such as Miles and Huberman (1994), advocate counting the frequencies of categories, which is a technique of quantitative content analysis. Similarly, the process of creating categories in quantitative content analysis is in fact a qualitative technique; statistical analysis often begins with a qualitative (visual) analysis of the data.

There is good reason for this mixing of methods of analysis. Earlier I reported on the arguments for using mixed-methods of data collection. The same arguments apply to using mixed-methods of data analysis: researchers can get a more holistic and valid view of a phenomenon by viewing and interpreting the phenomenon from different perspectives (Cresswell & Plano Clark, 2006; Greene & Caracelli, 1997; Johnson & Onwuegbuzie, 2004).

Questions to Consider

1. Of the four types of data analysis, what type are you most familiar and comfortable with? What advice would you give to a colleague intending to do that type of analysis for the first time?

2. In your own research community, is there an implicit hierarchy of types of data analysis? Is one type given more merit than others?

3. What do you think of the maxim from single-participant research that "if you cannot see an effect, it is probably not there"?

4. Gene V. Glass, the father of meta-analysis, once wrote "we need to stop thinking of ourselves as scientists testing grand theories, and face the fact that we are technicians collecting and collating information, often in quantitative forms" (2000, n.p.). As education researchers and data analysts, should we take Glass's advice and regard ourselves as "technicians collecting and collating information" or as "scientists testing grand theories"?

Chapter 6

Reporting Educational Technology Research

There is a prevalent misconception that the research process ends after data are collected. However, the truth is that reporting is one the most important stages of the research process. No matter how carefully controlled a study may have been, no matter the attention to detail in collecting and analyzing data, or no matter how large of a random sample was selected, a study is particularly vulnerable to bias at the reporting stage. For example, a common defect in many articles is that an author's interpretation and discussion of study outcomes is not congruent with what actually happened in the study. I have read countless educational technology studies in which the author carefully conducted an empirical study, but then completely ignored the study's empirical findings when formulating research conclusions. Other examples of introducing bias at the reporting stage include not reporting on relevant study complications, not providing information about study operations that could have affected study outcomes, and insufficiently operationalizing constructs and variables.

Besides being a stage of the research process that is particularly vulnerable to the introduction of bias, much of the potential utility and acceptability of the research is determined at the reporting stage. Despite any potential utility or relevance a study might have, poorly reported research is usually poorly regarded research. On this point Bruce Thompson, one of the major figures in social science research, writes,

> *I am convinced that poor writing will doom even the most significant manuscript (such as truthfully reporting the impending end of the world or viable cold fusion) to rejection, yet even a trivial report has a reasonable chance of being published somewhere if the manuscript is well-written. Spend whatever time it takes to communicate economically and with probity of style. (2001, p. 342)*

One way to help ensure that research is accepted by the research community is to follow the established conventions for reporting research. In this chapter I focus on those conventions. I also discuss alternative styles of reporting, and report on Huff's (1999) method for using exemplars.

Reporting on Quantitative Investigations

The reporting of empirical quantitative investigations in education research is somewhat formulaic. It consists of a title, an abstract, an introduction (including a brief review of the literature), a methods section, a results section, a discussion section, and a conclusion section.

Title

The title of a report should be as concise as possible, but still give readers a good sense of what the paper is about and spark the readers' interest. Some journals (e.g., *Educational Technology & Society*) limit the number of words in a title to ten.

Abstract

Abstracts in educational technology typically consist of a 120-300 word narrative summary of the article. According to the *Publication Manual of the American Psychological Association* (American Psychological Association, 2001) a narrative abstract of an empirical article should describe:

- *the problem under investigation, in one sentence if possible;*
- *the participants or subjects, specifying pertinent characteristics, such as number, type, age, sex......;*
- *the experimental method, including the apparatus, data-gathering procedures, complete test names......;*
- *the findings, including statistical significance levels; and*
- *the conclusions and the implications or applications. (p. 14)*

The abstracts for a review or theoretical article, methodological papers, or case studies differ slightly, but the underlying logic is basically the same. See American Psychological Association (2001).

There has been a call lately for structured or claim-based abstracts in education research (Kelly & Yin, 2007; Mosteller, Nave, & Miech, 2004). For use both as a description and as an exemplar, I will include the structured abstract from Mosteller, Nave, and Miech's article on structured abstracts in education research:

> Background*: Approximately 1,100 education journals collectively publish more than 20,000 education research articles each year. Under current practice, no systematic way exists to move the research findings from these studies into the hands of the millions of education practitioners and policymakers in the United States who might use them.*

> Purpose*: To help disseminate education research findings, we propose that education journals consider adopting a structured abstract, a structural innovation that focuses on the format of the article itself. The structured abstract would replace the paragraph-style narrative summary – typically either an American Psychological Association-style abstract or "editor's introduction" – now present at the beginning of many articles.*
>
> Intervention*: A structured abstract is a formal and compact summary of an article's main features and findings. As does a table or figure, it has a predictable structure that compresses information into a small space and can be read independently of the main body of the article. The structured abstract is longer and more detailed than the standard paragraph-style narrative summary. On the printed page, the structured abstract appears between the title and the main body of the article. It includes basic items applying to all articles (i.e., background, purpose, research design, and conclusions) and several additional items that apply to some articles but not to others (i.e., setting, population, intervention, data collection and analysis, and findings).*
>
> Research Design*: Analytic essay.*
>
> *Conclusions: The structured abstract is a viable and useful innovation to help practitioners and policymakers systematically access, assess, and communicate education studies and research findings. Relative to current practice, the structured abstract provides a more robust vehicle for disseminating research through traditional routes as well as through new channels made possible by emerging technologies. (p. 29)*

Introduction

The American Psychological Association's publication manual (American Psychological Association, 2001) suggests that an introduction should do three things: it should introduce the problem, it should review the previous research, and it should state the purpose and rationale for the study. The introduction should begin with a description of the problem and explain why the problem is important. It should also develop the background of a study by reviewing the literature. In an empirical article, the literature review does not necessarily have to be comprehensive; it just has to summarize the works relevant to the issue and emphasize the major findings and methodological issues in the previous research. Reviewing the findings of previous research will make it possible in the discussion to relate the author's finding with the findings from previous research. Also, the author should make it clear from the review of literature how the current research will make a contribution to the body of research on that subject (see American Education Research Association, 2006, for a list of ways that an article can make a contribution to the existing literature). Ideally, the literature review should serve as the basis for the purpose or rationale for the study.

There are many fine resources on how to conduct a literature review. Boote and Beile (2004, 2005) discuss the elements of quality literature reviews. Cooper (1984), Cooper and Hedges (1994), Glass, McGaw, and Smith (1981), and Lipsey and Wilson

(2001) are some examples of highly regarded books on conducting quantitatively oriented literature reviews.

In the problem introduction, the author shows that there is a problem. In the literature review section, the author demonstrates what has been done to help solve that problem and what still needs to be done (or what needs to be known) for that problem to be solved. In the rationale and purpose section, one should state how the current research will make a unique contribution to solving the problem. Near the end of the introduction it is also appropriate to state research questions or hypotheses and explain how those research questions or hypotheses are aligned with the previous research and aligned with the problem. The logic here is that answering the research questions will help make up for the lack of research (or development), which in turn, will help solve the problem.

Methods

The next section in an empirical paper in educational technology is the methods section. The methods section is customarily divided into subsections. If the research deals with human participants there should be a section on participants. That section should report on how the participants were selected, the number of participants (in each group, if there were groups), the number of participants who quit the study, and "report major demographic characteristics such as gender, age, and race/ethnicity and where possible and appropriate, characteristics such as socio-economic status [and] disability status…" (American Psychological Association, 2001, p. 18). In educational technology research, it would also be important to include information about participants' educational backgrounds, such as academic concentration, years and type of schooling, academic aptitude, previous experience with the intervention, and so on. In educationally oriented research, it might also be important to include information about the educational setting where the research occurred. One might then label this section as *Participants and Setting*. It is better to err on the side of providing too much information in this section than too little.

There might also be a subsection on the instruments used in the study. In educational technology research, one might give the specifications for the types of computers or software used.

Another subsection, which is mandatory, is the description of the procedure used. Depending on the type of study, this might include such information as what variables were examined, what measures were used, how data were collected, who collected the data, how long the data were collected, how much data were collected, how data were analyzed, how many groups there were, what experimental research design was used, how participants were assigned to groups, how the research assistants were trained, among many other details. In essence, one should provide enough detail here that another researcher could more or less replicate the study solely by reading the description.

Results

In the results section, the author summarizes the results in enough detail to justify the conclusions. In traditional, quantitative, education research papers, the text of the results section is reserved for the presentation of information and for pointing out important pieces of information. If presenting statistical results, useful guidelines are given in American Psychological Association (2001), Wilkinson and the APA Task Force on Statistical Inference (1996), and the American Education Research Association (2006). The interpretations, conclusions, and implications should be reserved for the discussion section.

Discussion

The discussion section typically begins by answering the research questions, based on the information presented in the results section. If there were complications in the research (e.g., if something did not go as planned), it should be reported in the beginning of the discussion section. A mistake often seen in the discussion section is that an author makes claims in the discussion section, but does not base the claims on the findings in the results section.

It is also customary to examine the similarities and differences between the current findings and the findings from the previous research. A good discussion will also explain what are the implications of the research for solving the problem mentioned in the introduction. In essence, the discussion section uses the information in the results section to relate back to the introduction – to the research questions, to the previous research, and, of course, to the original problem.

Conclusion – The Take-Home Message

One of my academic writing mentors gave me some great advice about writing and presenting conclusions. He argued that the average reader or audience member is only going to remember two or three ideas from a research study or presentation, at most. He suggested, then, that an author should conclude by picking out the two or three most important ideas and packaging them in what he called a "take-home message" – the message that the reader or audience will, with hope, remember after they get home. Basically, the take-home message is a statement, packaged in the clearest and most quotable way possible, that says (figuratively) to the reader or audience: "If you are only going to remember a few things about my research, I want you to remember these:... Here is why you should remember them:.... " The elements that work particularly well in a take-home message are statements of (a) what had been known already, (b) what more is known now as a result of the study, and (c) why anybody should care.

In the paragraphs above I summarized the conventions for reporting quantitative, empirical investigations when human participants are involved. For conventional quantitative papers when human participants are not involved, such as literature reviews, the sections might be slightly different, but the overall logic is the same: one should report on (a) the problem, (b) the previous research on the problem, (c) the approach to and rationale for solving the problem, (d) how the research was conducted, (e) the results that were found, and (f) the interpretation of the results as they relate back to the study's purpose.

For educational technology researchers who do quantitative research with human participants the most important resource for learning how to report results is probably the American Psychological Association's publication manual (American Psychological Association, 2001). Other helpful resources include and Alton-Lee (1998), American Education Research Association (2006), and Wilkinson and the APA Task Force on Statistical Inference (1999).

Common Defects in Quantitative Research Reporting

At least in quantitatively oriented social science articles, reviewers and editors tend to repeatedly find the same defects in manuscripts submitted for publication. Those defects, from American Psychological Association (2001), are listed below:

- *piecemeal publication, that is the separation of a single substantial report into a series of overlapping papers;*
- *the reporting of only a single correlation – even a significant correlation between two variables rarely has an interpretive value;*
- *the reporting of negative results without attention to power analysis;*
- *lack of congruence between a study's specific operations (including those related to the design and analysis) and the author's interpretation and discussion of the study's outcomes (e.g., failure to report the statistical test at the level being claimed);*
- *failure to report effect sizes;*
- *failure to build in needed controls, often for a subtle but important aspect of the study; and*
- *exhaustion of a problem – there is a difference between ongoing research that explores the limits of the generality of a research finding and the endless production of papers that report trivial changes in previous research. (p. 5)*

Reporting on Qualitative Investigations

Reporting qualitative research is considerably less formulaic than reporting quantitative research. Qualitative authors, in general, have much freedom to experiment with alternate forms of reporting and presenting research. Also, there is considerable variability between the reporting conventions (or lack of conventions) between each of the five major qualitative traditions.

Reporting Narrative Research

Narrative research reports come in many shapes and forms. Some authors employ a narrative chronology – following the life of an individual over time. Other authors begin with a critical event in the main character's life and then follow with narratives that precede and follow the critical event. Other authors of narrative research reports use a traditional five chapter format (introduction, literature review, methods, discussion, and conclusion). The narrative researcher is encouraged to employ the devices of biography and narration, such as transitions, metaphor, time-and-place shifts, and rich description of characters, settings, and events. Clandinin and Connelly (2000) and Czarniawska (2004) are key resources for learning how to write narrative reports.

Reporting Phenomenological Research

The structure of phenomenological reports, at least in Moustakas's (1994) brand of phenomenology, is more like a conventional research report than a narrative report. Moustaka (1994), as summarized in Cresswell (2007), gives the following suggestions for what sections are needed in a phenomenological report:

- *Chapter 1: Introduction and statement of topic and outline.*
- *Chapter 2: Review of the relevant literature.*
- *Chapter 3: Conceptual framework of the model.*
- *Chapter 4: Methodology.*
- *Chapter 5: Presentation of the data.*
- *Chapter 6: Summary, implications, and outcomes. (p. 187)*

Note that the sections listed above are more or less identical to the sections in a conventional quantitative research paper.

See Moustakas (1994) or Polkinghorne (1989) for a more detailed discussion of the phenomenological research report. See Grigsby and Megel (1995) for a good example of the use of figures or tables for illustrating the essence of a phenomenon.

Reporting Grounded Theory Research

Compared to Moustakas's (1994) brand of reporting phenomenological investigations, grounded theory reports have a much broader set of reporting guidelines. Strauss and Corbin (1990), as cited in Cresswell (2007), give the following broad guidelines for writing up grounded theory reports:

- *Develop a clear analytic story. This is to be provided in the selective coding phase of the study.*
- *Write on a conceptual level, with description kept secondary to concepts and the analytic story. This means that one finds little description of the phenomenon being studied and more analytic theory at an abstract level.*
- *Specify the relationships among categories. This is the theorizing part of the grounded theory found in axial coding when the researcher tells the story and advances propositions.*
- *Specify the variations and relevant conditions, consequences, and so forth for the relationships among categories. In a good theory, one finds variation and different conditions under which the theory holds. This means that the multiple perspectives or variations in each component of axial coding are developed fully. For example, the consequences in the theory are multiple and detailed. (p. 190)*

Besides Strauss and Corbin (1990), Chenitz and Swanson (1986) – especially May's (1986) chapter – is also a good resource for those writing up the grounded theory report.

Reporting Ethnographic Research

There is much written on the crafting of ethnographic reports. Van Maanen (1988) writes about four different types of *ethnographic tales*: realist tales in which the author takes on an objective, scientific voice and sticks to the factual description of culture; confessional tales in which the author concentrates on his or her own personal fieldwork experiences; and impressionistic tales, which are a combination of realism and confessionalism.

Wolcott (1994) provides further advice on writing up an ethnographic report. Wolcott suggests writing *description, analysis*, and *interpretation* sections. According to Wolcott, as cited in Creswell (2007, p. 193), the description might have the following characteristics: "chronological order, the researcher or narrator order, a progressive focusing, a critical or key event, plots and characters, groups in interaction, an analytical framework, and a story told through several perspectives." The analysis section might include "highlighting findings, displaying findings, reporting fieldwork procedures, identifying patterned regularities in the data, comparing the case with a known case, evaluating the information, contextualizing the information within a broader theoretical framework, critiquing the research process, and proposing a redesign of the study" (p. 193). For interpreting the study, "the researcher can extend the analysis, make inferences from the information, do as directed or suggested by gatekeepers, turn to theory, refocus the interpretation itself, connect with personal experience, analyze or interpret the interpretive process, or explore alternative formats" (p. 193).

In addition to van Maanen (1988) and Wolcott (1994), one might also find the following resources to be useful: Denzin (1989); Emerson, Fretz, and Shaw (1995); Hammersley and Atkinson (1995); and Richardson (1990).

Reporting Case Study Research

Like many types of qualitative reports, case studies can take on a variety of forms. Yin (2003) has written an excellent chapter in which he discusses the different formats and structures for composing a case study report. The four formats are the classical single case format, the multiple-case report that "contains both the individual case studies and some cross-case chapters" (p. 147), the multiple-case report that only contains a cross-case analysis, and the question-and-answer format.

Concerning the structure of case study reports, Yin describes four structures. They are listed below:

- the linear-analytic structure, which follows a conventional reporting structure;
- the comparative structure, in which the same case is compared under different conceptual models or alternative explanations;
- the chronological structure, in which the events of a case are reported as they occurred over time;
- the theory-building structure, in which evidence for a theory is presented in logical order;
- the suspense structure in which the answer is given in the first section and the supporting evidence and details are given later; and
- finally, an unsequenced structure, which Yin claims is most useful for purely descriptive case studies.

In addition to Yin (2003), those who are writing up a case study will find useful Lincoln and Guba (1985), Merriam (1988), or Stake (1995).

Alternative Styles of Reporting

Above, I have presented the guidelines for writing up a conventional scholarly report; however, scientific reporting does not necessarily have to be conventional. Torres, Preskill, and Piontek (2004) provide information on creative reporting strategies designed to bolster the utility of research results. Those strategies include involving stakeholders in the reporting process; incorporating modern technologies such as chat rooms and Internet forums; and using techniques such as photography, cartoons, poetry, and drama. Angela Thody, a proponent of alternative styles of reporting, adds to that list of alternative reporting strategies with examples such as:

- *readers' theatre (where researchers [act] their respondents' views);*
- *dance interpreting the emotions arising from findings;*
- *town meetings (researchers [report] their finding briefly as political speeches and then [invite] audience participation, assisted by mobile microphones);*
- *debates (six researchers [have] exactly three minutes to put their cases). (2006, p. 11)*

Thody (2006) adds that the adoption of alternative styles is a postmodern reaction to the positivistic, impersonal style of writing that currently dominates scientific reporting. It is also a reaction to what she calls the standardization of science reporting that results from publication manuals from organizations such as the American Psychological Association (APA) and the Modern Language Association (MLA). On this point she writes,

> *Why, I mused, in the USA and Canada, so often depicted as lands of freedom, is so little discretion allowed to, or taken by, highly intelligent academics on how to present their work? Why have APA guidelines for writing up psychology experiments been adopted so wholeheartedly by other disciplines?.... Even when there are no strictures, such as when academics present their research orally, why do so many academics still elect to 'read' their papers and eschew the livelier arts of demonstration and teaching? (p. 10)*

Using Exemplars

Whether one decides to use conventional or alternative means of reporting, the best piece of overall piece of advice I can give for writing up a report is to follow Huff's (1999) method for using exemplars – "document[s] already in the literature that accomplish the kind of task you are trying to accomplish *in an effective way*" (p. 55). Huff writes that exemplars "help authors define their purpose more clearly, can be used to structure and solve problems in writing, [and] are a springboard for innovating beyond the structures used by others" (p. 55). Huff's method for using exemplars is reported below:

> *Step 1: Specify Your Intent.... Describe what you hope to accomplish in your writing project in at least three different ways. Settle on one distinct description. Seek advice about this objective from advisers and your writing community. (pp. 57 – 58)*
>
> *Step 2: Choose Published Examples of Work with a Similar Objective.... Identify four or five exemplars of the kind of paper you want to write. If necessary, seek advice about the fit between these exemplars and your project, then settle on a final group of two or three works. (p. 59)*
>
> *Step 3: Examine the Structure and Tone of Successful Exemplars.... Outline in detail each exemplar in your final set, noting the proportion of text devoted to each topic. Also make notes about the "tone" used in your exemplar. Draw conclusions about the tacit rules for this kind of contribution to scholarly communication.... Identify the aspects of each exemplar that are particularly effective as well as any that are ineffective in communicating the author's purpose. (pp. 60 – 61)*
>
> *Step 4: Move... Beyond Your Exemplars... Use your analysis to identify where you are making a unique contribution to the literature and think about the ways to usefully depart from your exemplars to more effectively deliver the primary message of your paper. (p. 62)*

In Summary

In this chapter I have reported on the conventions for reporting both quantitative and qualitative reports, discussed alternative styles of reporting, and presented Huff's (1999) method of using exemplars. With hope the readers can utilize this information to create high-quality, relevant, and publishable papers in educational technology.

Questions to Consider

1. If you have had experience submitting your articles to journals or conferences, what lessons have you learned from the reviewers' feedback and the process itself?

2. There is a voice of growing discontent with the significant impact that the American Psychological Association's publication manual (2001) has on scientific reporting standards worldwide. In your opinion, is the American Psychological Association's influence advantageous because of its potential for the standardization of scientific reporting or is it an instance of American imperialism?

3. Gene V. Glass made the following statement about the deficiencies in the current state of education research:

    The conception of our work that held that "studies" are the basic, fundamental unit of a research program may be the single most counterproductive influence of all. This idea that we design a "study," and that a study culminates in the test of a hypothesis and that a hypothesis comes from a theory – this idea has done more to retard progress in educational research than any other single notion. Ask an educational researcher what he or she is up to, and they will reply that they are "doing a study," or "designing a study," or "writing up a study" for publication. Ask a physicist what's up and you'll never hear the word "study." (In fact, if one goes to http://xxx.lanl.gov where physicists archive their work, one will seldom see the word "study." Rather, physicists – the data gathering experimental ones – report data, all of it, that they have collected under conditions that they carefully described. They contrive interesting conditions that can be precisely described and then they report the resulting observations.) (2000, n.p)

    Why do you think that Glass is so against the idea of "the study" as the basic reporting unit in social science research? Do you agree or disagree with Glass, and why?

4. At the end of this chapter, I suggested that the best way to learn how to do academic reporting is to find exemplars and copy their format. Do you use exemplars in this way? If so, what exemplars have you used, for what purpose, and do you have any that are worth recommending to your colleagues or classmates?

Chapter 7

Conclusions about Methods Choice in Educational Technology Research and Development

I have provided much information about methods choice in educational technology research. In the first and second chapters, I identified several factors that need to be taken into consideration when making methods choices and discussed the research approaches amenable to educational technology research. In the third chapter, I described the methods choices that educational technology researchers tend to make. In the fourth, fifth, and sixth chapters I provided information about how to collect data, analyze data, and report findings. Therefore, for the concluding chapter I thought it might be useful to provide a summary of the process of making informed choices about methods choice in educational technology.

A key step in methods choice is arriving at appropriate research questions. However, arriving at appropriate research questions is often a difficult, iterative process. One needs to be clear, first of all, about what the purpose of the research is and what research has been done to meet the purpose. As one reviews the research, one probably will have to revise the research question many times to find a question (or development task) that advances the state of knowledge in some area. The review of literature is so important because one can only accidentally advance the state of knowledge in an area when one does not know what the state of the knowledge is.

After arriving at a question that can advance the state of the knowledge, the task is to take into consideration other factors that affect the choice of methods and make an initial decision about which approach to use. For example, it is important to consider whether the research question is best answered with a quantitative or qualitative approach (or both), the degree of generalizability that is needed, the degree of accuracy that is needed, and the degree of researcher and stakeholder participation in the investigation. Two critical factors are propriety and feasibility; if a method cannot be found that is both appropriate and feasible, then the question might need to be modified. After an overall approach is chosen, then one works out the details of data collection, data analysis, and reporting.

While the literature on a certain approach will provide much of the information on what methods of data collection, analysis, and reporting are possible, there are many variations within an approach that must be chosen and many details that should be worked out beforehand. These details might again affect the research questions.

So, in summary, there are many interrelated factors that come into play when choosing which methods to adopt: the purpose of the research, the state and type of previous research, the resources that are available, ethical restraints, among many other factors. The crux, then, of choosing the right methods is taking these factors into consideration and tailoring a balanced set of methods to the particular research situation. The art and science of methods choice begins with the case.

References

Alderson, J. C. 2000. Technology in testing: The present and the future. System, 28(4), 593 – 603.

Alexander, S. & Hedberg, J. G. 1994. Evaluating technology-based learning: Which model? In K. Beattie, C. McNaught, & S. Willis (Eds.), Interactive Multimedia in higher education: Designing for change in teaching and learning (233 – 244). Amsterdam: Elsevier.

Almstrum, V. L., Hazzan, O., Guzdzial, M. & Petre, M. 2005. Challenges to computer science education research. In Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education SIGCSE '05 (191 – 192). New York: ACM Press.

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.

American Psychological Association, 2001. Publication manual of the American Psychological Association (5th ed.). Washington, DC: American Psychological Association.

Atkinson, P., Coffey, A. & Delamont, S. 2003. Key themes in qualitative research: Continuities and changes. Walnut Creek, CA: Alta Mira.

Bannan-Ritland, B. 2003. The role of design in research: The integrated learning design framework. Educational Researcher, 32(1), 21 – 24.

Becher, T. 1989. Academic tribes and territories: Intellectual enquiry and the cultures of disciplines. Buckingham: Society for Research into Higher Education.

Bednarik, R. & Randolph, J. J. (in press). Levels of analysis of sparse eye-tracking data in visual attention studies of program comprehension. In R. I. Hammoud (Ed.), Passive eye monitoring: Algorithms, Applications and experiments. New York: Springer.

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 Education 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.

Boruch, R. 2007. Encouraging the flight of error: Ethical standards, evidence standards, and randomized trials. In G. Julnes & D. J. Rog (Eds.), Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. New Directions for Evaluation, no. 113 (55 – 74). San Francisco: Jossey-Bass.

Brand, S. 1999. The clock of the long now: Time and responsibility. London: Phoenix.

Burghar, J. & Turns, J. 1999a. A survey of research questions in Educational Technology Research & Development. Program for Educational Transformation through Technology (PETTT). University of Washington. Retrieved April 11, 2007, from http://depts.washington.edu/pettt/papers/reviews/3-ResearchQuestionsLitReview-EDTechR&D.doc

Burghar, J. & Turns, J. 1999b. A survey of research questions in Human-Computer Interaction. Program for Educational Transformation through Technology (PETTT). University of Washington. Retrieved April 11, 2007, from http://depts.washington.edu/pettt/papers/reviews/3-ResearchQuestionsLitReview-HCI.doc

Burghar, J. & Turns, J. 2000. A survey of research questions in Frontiers in Education. Program for Educational Transformation through Technology (PETTT). University of Washington. Retrieved April 11, 2007 from, http://depts.washington.edu/pettt/papers/reviews/3-ResearchQuestionsLitReview-FIE1997.doc

Caffarella, E. P. 1999. The major themes and trends in doctoral dissertation research in educational technology from 1977 through 1998. In Proceedings of selected research and development papers presented at the National Convention of the Association for Educational Communications and Technology (AECT). (ERIC Document Reproduction Service No. ED436178)

Campbell, D. T. 1969. Reforms as experiments. American Psychologist, 24, 409 – 429.

Charmaz, K. 2006. Constructing grounded theory. London: Sage.

Chen, J. F. 1997. Computer generated error feedback and writing process: A link. TESL-EJ, 2(3), A-1.

Chenitz, W. C. & Swanson, J. M. 1986. From practice to grounded theory: Qualitative research in nursing. Menlo Park, CA: Addison-Wesley.

Clandinin, D. J. & Connelly, F. M. 2000. Narrative inquiry: Experience and story in qualitative research. San Francisco: Jossey-Bass.

Clark, R. E. 1983. Reconsidering research on learning from media. Review of Educational Research, 53, 445 – 459.

Clark, R. E. & Snow, R. E. 1975. Alternative designs for instructional technology research. AV Communications Review, 23, 373 – 394.

Cohen, J. & Cohen, P. (1983). Multiple regression/correlation for the social sciences (2nd ed.). Mahwah, NJ: Lawrence Earlbaum.

Comroe, J. H. & Drips, R. D. 1976. Scientific basis for the support of biomedical science. Science, 192, 105 – 111.

Constantine, L. L. & Lockwood, L. A. 1999. Software for use: A practical guide to the models and methods of usage-centered design. Boston: Addison Wesley Professional.

Cooper, H. M. 1984. The integrative research review: A systematic approach. Applied social research methods series (Vol. 2). Beverly Hills, CA: Sage.

Cooper, H. & Hedges, L. V. (Eds.) 1994. The handbook of research synthesis. New York: Sage.

Creswell, J. W. 1994. Research design: Qualitative and quantitative approaches. Thousand Oaks, CA: Sage.

Creswell, J. W. 2007. Qualitative inquiry & research design: Choosing among five approaches. (2nd ed.) Thousand Oaks, CA: Sage.

Creswell, J. W. & Plano Clark, V. L. 2006. Designing and conducting mixed methods research. Thousand Oaks, CA: Sage.

Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., et al. 1980. Towards reform of program evaluation. San Francisco: Jossey-Bass.

Cronjé, J. C. 2006. Pretoria to Khartoum – how we taught an Internet-supported Masters' programme across national, religious, cultural and linguistic barriers. Educational Technology & Society, 9(1), 276 – 288.

Czarniawska, B. 2004. Narratives in social science research. London: Sage.

Datta, L-E. 1994. Paradigm wars: A basis for peaceful coexistence and beyond. In C.S. Reichardt and S. F. Rallis (Eds.), The quantitative-qualitative debate: New perspectives. New Directions for Program Evaluation, no. 61 (pp. 53 – 70). San Francisco: Jossey-Bass.

De Leeuw, E. D. 1993. Data quality in mail, telephone and face-to-face surveys. Amsterdam: TT-Publikates. (ERIC Document Reproduction Service No. ED374136).

De Leeuw, E. D. & Nicholls, W. 1996. Technological innovations in data collection. Acceptance, data quality and costs. Sociological Research Online, 1(4).

Denzin, N. K. 1989. Interpretive biography. Newbury Park, CA: Sage.

Design-Based Research Collective. (n.d.) The Design-Based Research Collective website. Retrieved May 9, 2007, from http://www.designbasedresearch.org/

Design-Based Research Collective, 2003. Design-based research: An emerging paradigm for educational inquiry. Educational Researcher, 32(1), 5 – 8.

Dick, W. & Carey, L. 1990. The systematic design of instruction. Glenview, IL: Scott, Foresman/Little.

Dick, W. & Dick, W. D. 1989. Analytical and empirical comparisons of the Journal of Instructional Development and Educational Communication and Technology Journal. Educational Technology Research & Development, 37(1), 81 – 87.

Dilman, D. A. 1978. Mail and telephone surveys: The total design method. New York: Wiley.

Driscoll, M. P. & Dick, W. 1999. New research paradigms in instructional technology. An inquiry. Educational Technology Research & Development, 47(2), 7 – 18.

Durland, M. M. & Fredericks, K. A. (Eds.) 2006. Social network analysis. New Directions for Evaluation, no. 107. San Francisco: Jossey-Bass.

Educational Testing Service. (n.d.) The Test Collection. Retrieved June 14, 2007, from http://www.ets.org/testcoll/

Emerson, R. M., Fretz, R. I. & Shaw, L. L. 1995. Writing ethnographic fieldnotes. Chicago: University of Chicago Press.

Fink, A. 2002. The survey kit (2nd ed.). Thousand Oaks, CA: Sage.

Frechtling, J., Frierson, H., Hood, S. & Hughes, G. 2002. The user friendly handbook for project evaluation. NSF02-057. Arlington, VA: NSF. Retrieved May 10, 2005, from http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

Frechtling, J. & Sharp, L. 1997. The user friendly handbook for mixed-method evaluations. NSF 97-153. Arlington, VA: NSF. Retrieved May 10, 2007, from http://www.nsf.gov/pubs/1997/nsf97153/start.htm

Frechtling, J., Stevens, F., Lawrenz, F. & Sharp, L. 1993. The User Friendly handbook for project evaluation: Science, mathematics and technology education. NSF 93-152. Arlington, VA: NSF.

Gall, M. D., Borg, W. R. & Gall, J. P. 1996. Educational research: An introduction (6th ed.). White Plains, NY: Longman.

Gay, G. & Bennington, T. L. (Eds.) 1999. Information technologies in evaluation: Social, moral, epistemological, and practical implications. New Directions for Evaluation, no. 84. San Francisco: Jossey-Bass.

Geertz, C. 1973. The interpretation of cultures. New York: Basic.

Glass, G. V. 2000, January. Meta-analysis at 25. Retrieved August 21, 2007, from http://glass.ed.asu.edu/gene/papers/meta25.html

Glass, G. V., McGaw, B. & Smith, M. L. 1981. Meta-analysis in social research. Beverly Hills, CA: Sage.

Glatthorn, A., Bragaw, D., Dawkins, K. & Parker, J. 1998. Performance assessment and standards-based curricula: The achievement cycle. Larchmont, NY: Eye on Education.

Gorard, S. & Taylor, C. 2004. Combining methods in educational and social research. Berkshire, England: Open University Press.

Greene, J. C. & Caracelli, V. J. (Eds.). 1997. Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. New Directions for Program Evaluation, no. 74. San Francisco: Jossey-Bass.

Greene, J. C., Lipsey, M. W., Schwandt, T. A., Smith, N. L. & Tharp, R. G. 2007. Methods choices are contextual, contingent, and political. In G. Julnes & D. J. Rog (Eds.), Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. New Directions for Evaluation, no. 113 (111 – 128). San Francisco: Jossey-Bass.

Grigsby, K. A. & Megel, M. E. 1995. Caring experiences of nurse educators. Journal of Nursing Research, 34, 411 – 418.

Guba, E. G. 1990. The alternative paradigm dialog. In E. G. Guba (Ed.), The paradigm dialog (17 – 30). Thousand Oaks, CA: Sage.

Hammersley, M. & Atkinson, P. 1995. Ethnography: Principles in practice (2nd ed.). New York: Routledge.

Herman, J. L. (Ed.) 1987. Program evaluation kit (2nd ed.) Newbury Park, CA: Sage.

Higgins, N., Sullivan, H., Harper-Marinick, M. & Lopez, C. 1999. Perspectives on educational technology research and development. Educational Technology Research & Development, 37(1), 7 – 17.

Huck, S. W. 2007. Reading statistics and research (5th ed.). Boston: Allyn & Bacon.

Huff, A. S. 1999. Writing for scholarly communication. Thousand Oaks, CA: Sage.

Isaac, S. & Michael, W. B. 1990. Handbook in research and evaluation: A collection of principles, methods and strategies useful in the planning, design and evaluation of studies in education and the behavioral sciences. San Diego, CA: EdITs.

Jarvinen, P. H. 2000. Research questions guiding selection of an appropriate research method. In H. R. Hansen, M. Bichler, H. Mahrer (Eds.), Proceedings of the 8th European Conference on Information Systems (124 – 131). Vienna: Wirtschaftsunivsitat Wien.

Johnson, B. R. & Onwuegbuzie, A. J. 2004. Mixed methods research: A paradigm whose time has come. Educational Researcher, 33(7), 14 – 26.

Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards (2nd ed.). Thousand Oaks, CA: Sage.

Julnes, G. & Rog, D. J. 2007a. Editor's notes. In G. Julnes & D. J. Rog (Eds.), Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. New Directions for Evaluation, no. 113 (1 – 12). San Francisco: Jossey-Bass.

Julnes, G. & Rog, D. J. (Eds.) 2007b. Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation. New Directions for Evaluation, no. 113. San Francisco: Jossey-Bass.

Kazdin, A. E. 1982. Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press.

Kelly, A. E. (Ed.) 2003. Theme issue: The role of design in educational research. Educational Researcher, 32(1).

Kelly, A. E. & Yin, R. K. 2007. Strengthening structured abstracts for education research: The need for claim-based structured abstracts. Educational researcher, 33, 133 – 138.

Kiesler, S. & Sproull, L. S. 1988. Response effects in the electronic survey. Public Opinion Quarterly, 50, 402 – 413.

Klein, J. D. 1997. ETR&D development: An analysis of content and survey of future direction. Educational Technology Research & Development, 45(3), 57 – 62.

Krippendorf, K. 1980. Content analysis. An introduction to its methodology. Beverly Hills, CA: Sage.

Kvale, S. 1996. Interviews: An introduction to qualitative research interviewing. Thousand Oaks, CA: Sage.

Laakso, M-J., Salakoski, T., Korhonen, A. & Malmi, L. 2004. Automatic assessment of exercises for algorithms and data structures – a case study with TRAKLA2. In A. Korhonen and L. Malmi (Eds.), Kolin Kolistelut – Koli Calling 2004, Proceedings of the Fourth Annual Finnish Annual Finnish Baltic Sea Conference on Computer Science Education (28 – 36). Finland: Helsinki University of Technology.

LeCompte, M. D. & Schensul, J. J. 1999. Designing and conducting ethnographic research. Walnut Creek, CA: AltaMira.

Lincoln, Y. S. & Guba, E. G. 1985. Naturalistic inquiry. Beverly Hills, CA: Sage.

Lipsey, M. W. & Wilson, D. B. 2001. Practical meta-analysis. Applied social research methods series (Vol. 49). Thousand Oaks, CA: Sage.

Madison, D. S. 2005. Critical ethnography: Methods, ethics, and performance. Thousand Oaks, CA: Sage.

Mark, M. M., Henry, G. T. & Julnes, G. 2000. Evaluation: An integrated framework for understanding, guiding, and improving policies and programs. San Francisco: Jossey-Bass.

May, K. A. 1986. Writing and evaluating the grounded theory research report. In W. C. Chenitz & J. M. Swanson (Eds.), From practice to grounded theory (146 – 154). Menlo Park, CA: Addison-Wesley.

Merriam, S. B. 2001. Qualitative research and case study applications in education. San Francisco: Jossey-Bass.

Mosteller, F., Nave B. & Miech, E.J. 2004. Why we need a structured abstract in education research. Educational Researcher, *33*(1), 29 – 34.

Miles, M. B. & Huberman, A. M. 1994. Qualitative data analysis: An expanded sourcebook. Thousand Oaks, CA: Sage.

Mohr, L. B. 1999. The qualitative method of impact analysis. American Journal of Evaluation, 20(1), 69 – 84.

Moustakas, C. 1994. Phenomenological research methods. Thousand Oaks, CA: Sage.

Neuendorf, K. A. 2002. The content analysis guidebook. Thousand Oaks, CA: Sage.

Office of Research Services, 2007. Guidelines for MS Plan A thesis and doctoral dissertation proposals. Utah State University, retrieved July 30, 2007, from http://www.cehs.usu.edu/ors/PDF%20files/PLANA.pdf

Osorio, N. L. 1998. Content Analysis of Web Pages for Distance Learning Engineering Programs. In H. Maurer & R. G. Olson (Eds,), Proceedings of WebNet 98: World Conference of the WWW, Internet, & Intranet, (1 – 8), Orlando, FL: WebNet.

Patton, M. Q. 1990. Qualitative evaluation and research methods. Newbury Park, CA: Sage.

Phipps, R. & Merisotis, J. 1999, April. What's the difference: A review of the contemporary research on the effectiveness of distance learning in higher education. The Institute for higher education policy. Retrieved July 27, 2006, from http://www.ihep.org/Pubs/PDF/Difference.pdf

Polkinghorne, D. E. 1989. Phenomenological research methods. In R. S. Valle & S. Halling (Eds.), Existential-phenomenological perspectives in psychology (41 – 60). New York: Plenum Press.

Potter, J. & Wetherell, M. 1987. Discourse and Social Psychology. London: Sage.

Preskill, H. & Torres, R. T. 1999. Evaluative inquiry for learning in organizations. Thousand Oaks, CA: Sage.

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 April 11, 2007, from http://www.archive.org/details/randolph_dissertation

Randolph, J. J. 2007b. A methodological review of the program evaluation in K-12 computer science education research. Manuscript submitted for publication.

Randolph, J. J. 2007c. What's the difference, still? A follow-up critique of the contemporary quantitative research methodology in distance learning. Informatics in Education, 6(1), 179-188.

Randolph, J. J., Bednarik, R., Silander, P., Lopez-Gonzalez, J., Myller, N. & Sutinen, E. 2005. A critical review of research methodologies reported in the full-papers of ICALT 2004. In Kinshuk, C-K Looi, E. Sutinen, D. Sampson, I. Aedo, L. Uden, & E. Kähkönen (Eds.), Proceedings of the Fifth International Conference on Advanced Learning Technologies, (10 – 14). Los Alamitos, CA: IEEE Computer Society.

Randolph, J. J., Virnes, M., Jormainen, I. & Eronen, P. J. 2006. The effects of a computer-assisted interview tool on data quality. Educational Technology & Society, 9(3), 195 – 205.

Reason, P. & Bradbury, H. (Eds.) 2001. Handbook of action research: Participative inquiry and practice. Sage: Thousand Oaks, CA.

Reeves, T. C. 1995. Questioning the questions of instructional technology research. Retrieved October 19, 2004, from http://www.hbg.ps.edu/bsed/intro/docs/dean/

Richardson, L. 1994. Writing strategies: Reaching diverse audiences. Newbury Park, CA: Sage.

Rieger, R. H. & Sturgill, A. 1999. Evaluating on-line learning environments: Tools for observing users and gathering feedback. In G. Gay and T. L. Bennington (Eds.), Information technologies in evaluation: Social, moral, epistemological, and practical implications. New Directions for Evaluation, no. 84 (45 – 58). San Francisco: Jossey Bass.

Rogers, E. M. 1995. Diffusion of innovations. New York: Free Press.

Romesburg, H. C. 2004. Cluster analysis for researchers. North Carolina: Lulu Press.

Ross, S. M. & Morrison, G. R. 2004. Experimental research methods. In D. J. Jonassen (Ed.) Handbook of research on educational communications and technology (2nd ed., 1021 – 1043.) Mahwah, NJ: Erlbaum.

Rubin, H. J. & Rubin, S. I. 2004. Qualitative interviewing: The art of hearing data. Thousand Oaks, CA: Sage.

Rudner, L. & Gagne, P. 2001. An overview of three approaches to scoring written essays by computer. Practical Assessment, Research & Evaluation, 7(26).

Saris, W. E. 1989. A technological revolution in data collection. Quality and Quantity, 23, 333 – 350.

Saris, W. R. 1991. Computer assisted interviewing (Quantitative applications in the social sciences, no. 80). Newbury Park: Sage.

Scriven, M. 1976. Maximizing the power of causal investigation. In G. V. Glass (Ed.), Evaluation studies review annual (Vol. 1, 101 – 118). Newbury Park, CA: Sage.

Scriven, M. S. (Ed.) 1993. Hard-won lessons in program evaluation. New Directions for Evalution, no 58. San Francisco: Jossey-Bass.

Shadish, W. R., Cook, T. D. & Campbell, D. T. 2002. Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin.

Silka, L. 1989. Intuitive judgments of chance. New York: Springer-Verlag.

Spector, P. 1992. Summated rating scale construction: An introduction. Sage University Paper Series: Quantitative Applications in the Social Sciences, no. 82. New Bury Park, CA: Sage.

Stake, R. 1995. The art of case study research. Thousand Oaks, CA: Sage.

Stevens, J. P. 1999. Intermediate statistics: A modern approach (2nd ed.). Mahwah, NJ: Lawrence Earlbaum.

Stewart, D. W., Shamdasani, P. N. & Rook, D. W. 2006. Focus groups: Theory and practice (2nd ed.). Thousand Oaks, CA: Sage.

Stokes, D. E. 1997. Pasteur's Quadrant: basic science and technological innovation. Washington, D.C.: The Brookings Institution.

Strauss, A. & Corbin, J. 1990. Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park, CA: Sage.

Sutinen, E., Virmajoki-Tyrväinen, M. & Virnes, M. 2005. Physical Learning Objects Can Improve Social Skills in Special Education. In A. Antikainen (Ed.), Transforming a Learning Society: The case of Finland. Explorationen. Studien zur Erziehungswissenschaft (117 – 130). Bern: Peter Lang.

Tashakkori, A. & Teddlie, C. 2002. Handbook of mixed methods in social & behavioral research. Thousand Oaks, CA: Sage.

Texas Center for Educational Technology (n.d.) Attitudes towards computers. Retrieved June 14, 2007, from http://www.tcet.unt.edu/pubs/attcomp.htm

Thody, A. 2006. Writing and presenting research. Thousand Oaks, CA: Sage.

Thompson, B. 2001. Publishing your research results: Some suggestions and counsel. Journal of Counseling & Development, 73, 342 – 345.

Thompson, B. 2004. Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association.

Thorndike, R. M. 2004. Measurement and evaluation in psychology and education (7th ed.). New York: Prentice Hall.

Thorndike, R. M. & Dinnel, D. 2000. Basic statistics for the behavioral sciences. New York: Prentice Hall.

Torres, R. T., Preskill, H. S. & Piontek, M. E. 2004. Evaluation strategies for communicating and reporting: Enhancing learning in organizations (2nd ed.). Thousand Oaks, CA: Sage.

Ulrich, K. T. & Eppinger, S. D. 2000. Product design and development. New York: McGraw-Hill.

U.S. Department of Education, 2003. Scientifically based evaluation methods. Federal Register, 68, 62445 – 62447.

van Maanen, J. 1988. Tales of the field: On writing ethnography. Chicago: University of Chicago Press.

van Manene, M. 1990. Researching lived experience: Human science for an action sensitive pedagogy. Albany: State University of New York Press.

Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., Mislevy, R. J., Steinberg, L., et al. 2000. Computer adaptive testing: A primer. Mahwah, NJ: Lawrence Earlbaum.

Weiss, C. H. 1998. Evaluation: Methods for studying programs and policies (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Weitzman, E. A. & Miles, M. B. 1995. Computer programs for qualitative data analysis. Thousand Oaks, CA: Sage.

Wiggins, G. P. 1993. Assessing student performance: Exploring the purpose and limits of testing. San Francisco: Jossey:Bass.

Wilkinson, L. & the APA Task Force on Statistical Inference, 1999. Statistical methods in psychology journals: Guidelines and explanations [Electronic version]. American Psychologist, 54, 594 – 604.

Williamson, A., Nodder, C. & Baker, P. 2001. A critique of educational technology research in New Zealand since 1996. Retrieved July 25, 2007, from http://www.ascilite.org.au/conferences/melbourne01/pdf/papers/williamsona.pdf

Wolcott, H. F. 1994. Transforming qualitative data: Description, analysis, and interpretation. Thousand Oaks, CA: Sage.

Wolcott, H. F. 1999. Ethnography: A way of seeing. Walnut Creek, CA: AltaMira.

Yin, R. K. 2003. Case study research: Design and methods (3rd ed.). Thousand Oaks, CA: Sage.

Appendix

Key Questions in Educational Technology Methods Choice

In the preceding chapters, I identified several factors important in educational technology methods choice. I used those factors to create a list of key questions in methods choice.

This list can be used for several purposes. First it can be used to introduce students to the factors that are important to consider when planning educational technology research. In that capacity, it might serve as a classroom teaching aid or as an aid for supervising student researchers in the early stages of planning their theses or dissertations.

Second, it can help beginning researchers reflect on, understand, and clarify their own research plans and come to understand how those research plans relate to other research done in the field of educational technology. Students will come to understand how their research fits into the overall picture of the research being done in their field.

Third, the list of key questions in methods choice can be used to create a research profile and identify others with similar profiles. For example, in a research methods course, students could create a research profile by elucidating their answers to the list of key questions and try to find others who have similar research profiles. Or, in the future, a research profiler based on the key questions might be automated so that researchers with similar profiles can be matched (e.g., researchmatch.com).

Fourth, the research profile, with some modifications, can be used as a framework for doing research about research. For example, one could compare the research profiles from one genre of research to the profiles in another genre. And, finally, the research profile can be used as a guide to methods choice.

Key Questions in Educational Technology Methods Choice

1. What is the research problem that you plan to resolve?
2. Is there a social problem related to your research problem? If so, what is it?
3. What is the primary purpose of your research?
    a. To develop an intervention.
    b. To answer questions important to local stakeholders.
    c. To answer questions important to the scientific community.

4. What type of research contribution do you intend to make?
    a. Contribute to an already established theory or line of empirical research.
    b. Help establish a new theory.
    c. Meet a practical need.
    d. Make up for a lack of needed information about a topic.
    e. Other.

5. If you are investigating a phenomenon, what is the state of theoretical knowledge about the phenomenon?
    a. There are no established theories.
    b. There are theories, but they are not yet substantiated.
    c. There are substantiated theories, but new theories need to be developed.
    d. There are substantiated theories, and they are sufficient.

6. If you are investigating a phenomenon, what is the state of empirical knowledge about the phenomenon?
    a. The important variables or the elements of a phenomenon have been identified.

b. The associations between those elements or variables have been substantiated.

c. The causal mechanisms regarding the phenomenon are clear.

7. Which of the previous studies are related to your research and how are they related? (A good way to answer this question is to create a research map.)

8. Who are the major researchers in your field?

9. What research methods were used in the previous research? (For example, you might answer this question by making a table in which you describe the following characteristics of the previous studies, as applicable: the research approaches, the methods of data collection, the methods of analysis, the variables examined, the settings involved, the participants involved, or other salient characteristics.)

10. What is your general research question and how does it relate to the research questions asked in the previous research?

11. What are the sub-questions that unpack the general research questions? Are they procedural or structural or are they another type of research sub-question?

12. Which of your research questions are meant to be answered by examining the knowledge-base and which are meant to be examined empirically?

13. What category of educational technology research question does your general research question fall into?

    a. Questions about theories and the practice of educational technology.

    b. Questions about research and development methods.

    c. Questions about technology implementation.

    d. Questions about the effectiveness of an intervention.

    e. Questions about the factors that moderate the effectiveness of an intervention.

    f. Other.

14. Which of the following research acts are implied in your research question?

    a. Identification.

    b. Description.

    c. Comparison.

d. Correlation.

e. Experimentation (causal description).

f. Causal explanation.

15. What family of research approaches do you intend to use, and why?

a. The quantitative family.

b. The qualitative family.

c. Mixed-methods.

16. What research approach, or approaches do you intend to use, and why? Do your research approaches match up with the research acts in Question 12?

a. Survey research.

b. Causal-comparative research.

c. Correlational research.

d. Experimental research.

e. Narrative research.

f. Phenomenological research.

g. Ethnographic research.

h. Case study research.

i. Grounded theory research.

j. Other.

17. To what degree do you intend for your research to generalize across participants, interventions (or phenomena), outcomes, and settings?

18. To what degree do you intend to do exploratory or confirmatory research?

19. To what degree will you involve stakeholders in the research process?

20. To what degree will you (the researcher) be involved in the research setting or involved in the phenomenon being investigated?

21. How accurate do your findings have to be? (e.g., How many participants will you need? How many pages of transcripts do you intend to get?)

22. Who are the possible audiences for your findings and how will you disseminate your findings to them in a way to ensure that your results are timely, informative, and influential?

23. What safeguards are in place to ensure that your research is ethical?

24. What are the time and resources necessary to carry out your proposed research? (Create a budget and timeline, including estimated work hours.)

25. What is your argument why the benefits that will come about as a result of your research are worth the time and resources necessary to carry out the study?

Over the past thirty years, there has been much dialogue, and debate, about the conduct of educational technology research and development. In this brief volume, Justus Randolph helps clarify that dialogue by theoretically and empirically charting the research methods used in the field and provides much practical information on how to conduct educational technology research. Within this text, readers can expect to find answers to the following questions:

- What are the methodological factors that need to be taken into consideration when designing and conducting educational technology research?

- What types of research questions do educational technology researchers tend to ask?

- How do educational technology researchers tend to conduct research? – What approaches do they use? What variables do they examine? What types of measures do they use? How do they report their research?

- How can the state of educational technology research be improved?

In addition to answering the questions above, the author, a research methodologist, provides practical information on how to conduct educational technology research – from formulating research questions, to collecting and analyzing data, to writing up the research report – in each of the major quantitative and qualitative traditions. Unlike other books of this kind, the author address some of research approaches used less commonly in educational technology research, but which, nonetheless, have much potential for creating new insights about educational phenomena – approaches such as single-participant research, quantitative content analysis, ethnography, narrative research, phenomenology, and others.

Multidisciplinary Methods in Educational Technology Research and Development is an excellent text for educational technology research methods courses, a useful guide for those conducting (or supervising) research, and a rich source of empirical information on the art and science of educational technology research.

HAMK
UNIVERSITY OF APPLIED SCIENCES

ISBN 978-951-784-456-7 (PDF)
ISSN 1795-424X
HAMKin e-julkaisuja 1/2008