Clipping and Coding Audio Files: A Research Method to Enable Participant Voice

Susan Crichton and Elizabeth Childs

Susan Crichton, PhD, Assistant Professor, Faculty of Education, University of Calgary, Calgary, Canada

Elizabeth Childs, PhD, Assistant Professor, Faculty of Medicine, University of Calgary, Calgary, Canada

Abstract: *Qualitative researchers have long used ethnographic methods to make sense of complex human activities and experiences. Their blessing is that through them, researchers can collect a wealth of raw data. Their challenge is that they require the researcher to find patterns and organize the various themes and concepts that emerge during the analysis stage into a coherent narrative that a reader can follow. In this article, the authors introduce a technology-enhanced data collection and analysis method based on clipped audio files. They suggest not only that the use of appropriate software and hardware can help in this process but, in fact, that their use can honor the participants' voices, retaining the original three-dimensional recording well past the data collection stage.*

Keywords: *computers, interviews, transcripts, ethnography*

Introduction

Qualitative researchers have been attempting to make sense of the world around them since armchair anthropologists and sociologists[1] left the confines of their familiar environments and ventured into the field (Crichton, 1997). In this article, we suggest that today, little has changed in terms of the intent of that work; however, the tools and methods used have evolved in such a way that the voice of the participants under study has the potential to be stronger and more authentic when digital audio recordings are clipped and coded, as discussed in this article, rather than directly transcribed.

Researchers (Crichton & Kinash, 2003) have suggested, "Ethnography is a qualitative field of research intended to construct in-depth depictions of the every day life events of people through active researcher participation and engagement" (p. 102). One of the challenges in this type of work is the sheer quantity of data captured and the need to develop an elastic yet rigorous structure in which to organize and analyze it. This structure allows the researcher to share findings, typically presented in a narrative form, in such a way that a reader, without the same firsthand experiences, can make personal sense from the rich description. It is this narrative that "allow[s] readers to understand fully the research setting and the thoughts of the people represented . . . [stopping] short, however, of becoming trivial and mundane" (Genzuk, 2004, p. 10).

A further challenge for the researcher is determining a balance between the essential description required to set the context and the critical analysis and interpretation that is necessary to help the reader come to an understanding of the findings. When that balance is achieved, the reader is able to interpret and understand the work and make relevant links that might extend the work and generalize the findings. However, as Genzuk (2004) has cautioned, "Endless description becomes its own muddle. [The] . . . purpose of analysis is to organize the description in a way that makes it manageable" (p. 10).

Typically, research findings that have been reached using qualitative methods will include thick descriptions of the experiences, contexts, and general environment of the site, individuals, and/or phenomenon under study. These descriptions are critical for the reader to understand the details of what has happened and the various viewpoints of the participants. They help to create a holistic picture, so that the findings do not lose their credibility and impact. Traditionally, these findings have been reached through the analysis of transcribed data or using qualitative analysis software such as NUD*IST. Our purpose in this article is to offer an alternative method for collecting and analyzing qualitative data: digital audio files.

Ethnography as process

Hammersley (1990) explained ethnography as social research that involves the study of people within their everyday situations rather than in experimental conditions. It is an attempt to collect data from a range of sources, including formal and informal conversations and observations as well as tangible and intangible artifacts. Although it might appear that the approach to the data collection is unstructured and at times chaotic and eclectic, it does not suggest that the collection of data has not been considered and planned. Ethnography simply encourages the ongoing collection of raw data from the widest and richest range of possible sources, continuing the collection until the field is either saturated or the data become repetitive. Merriam (1998) has suggested that the "emergence of regularities" serves as the indicator that a sufficient amount of data had been collected, noting further that "data collection in a case study is a recursive, interactive process in which engaging in one strategy incorporates or may lead to subsequent sources of data." (p. 134)

Put simply, ethnography is consistent with the way in which many of us make sense of events in our daily lives. Ethnography "literally means a portrait of a people . . . [as it is] a written description of a particular culture—the customs, beliefs, and behavior—based on information collected through fieldwork" (Harris & Johnson, 2000). Typically, data are collected from structured or semistructured interviews, formal and informal observation, and the analysis of various documents developed by or found at the site under study. From these data, the researcher is able to analyze, interpret, and share direct quotations,

thick descriptions, and selected excerpts (Hammersley, 1990), merging them into a rich narrative that forms a complex narrative quilt that covers the essential aspects of the study.

This merging process begins with the analysis of the various data. The researcher starts by labeling and sorting the various items into a type of order that allows him or her to make sense of what is there and begin to group items into categories. As the grouping continues, it begins to give the researcher a sense of what is there, what is missing, and whether the data-gathering phase is nearing completion. It is during this process that patterns begin to emerge and themes arising from the researcher's previous work or literature review are supported or rejected.

Critical to the analysis is a full disclosure of the role of the researcher in the study. It is important that the reader understand the researcher's background and personal experience, as it is through her or his eyes that the reader will come to understand the work. Therefore, declaring whether the researcher has been simply an observer or a participant observer is important, as access to the field must be seen as a privilege, not a right, and the participants and the events must be treated with respect. As an observer, the researcher might have no previous knowledge of the context or the participants, whereas a participant observer might have an intimate understanding and an ongoing relationship. Genzuk (2004) has cautioned, "we cannot assume that we already know others' perspectives, even in our own society, because particular groups and individuals develop distinctive worldviews" (p. 4). Therefore, the researcher has an obligation to declare her or his role and background, so that readers can make the critical inferences that might allow them to generalize the findings to their own understandings and circumstances.

Because ethnography supports inquiry into the experiences under study, it is critical that the researcher not approach the field with an explicit set of points to prove. "It is argued that if one approaches a phenomenon with a set of hypotheses one may fail to discover the true nature of that phenomenon, being blinded by the assumptions built into the hypotheses" (Genzuk, 2004, p. 4). Simply put, the researcher will be able to find only what he or she knew in advance to look for.

Considering digital tools

As stated earlier, a challenge in ethnography, case study, and many other types of qualitative research is managing the quantity of data collected. As the data collection process draws its strength from accumulating raw data and making sense of it without a set of preconceived hypotheses, the analysis process is labor intensive and time consuming. Researchers wishing to use the ethnographic approach usually do so because they wish to honor and record the participants' lived experiences in situ, allowing them to speak for themselves without misinterpretation.

It is our position that the use of digital video allows us to be the unobtrusive observers suggested in the literature (Kellehear, 1997). The method we suggest in this article allows us to record our participants, analyze their words and actions, clip the relevant segments, and organize those segments into a series of frames and codes (Goffman, 1959, 1974), keeping the images and/or voice of the participants intact for as long as possible. This method allows the researcher to hear and see the gestures, intonation, passion, pauses, and inflections throughout the analysis process. It reduces the impact that the transcription process has on the content, given that often the transcription is not done by the principal researcher because of time or cost considerations and that the transcription process itself flattens the potentially rich, three-dimensional quality of the original footage into a two-dimensional text format. Furthermore, it recognizes that "researchers have traditionally spent large proportions of their budgets on interview transcription, [which, in turn, has influenced] the number and extent of interviews based on cost (time and money) projections" (Crichton & Kinash, 2003, p. 104).

We have described the method of digital audiorecording, clipping, and coding in the three case studies below. This method involves the use of digital video- and film-editing software to support the development of frames and codes for data analysis. Two of the cases report on interviews that were recorded using a digital video camera. The interview files are then placed in audio editing software (iMOVIE or Final Cut Express), where they are edited initially by the researcher for extraneous content, such as opening comments, interruptions, discussions that are informal and not part of the research, and so on. After this initial sort, the files are exported as QuickTime movies, which are analyzed and chunked into appropriate frames and codes. These clipped audio files are then placed, according to the

emerging themes, in a Microsoft Excel spreadsheet for further sorting and analysis. In the cases reported in this article, only the audio portion of the interview was analyzed; we did not set out to capture video images.

We had originally hoped that Final Cut Express would allow us to sort and organize our clips using the database within the software, but we could not get this function to work. Therefore, we used Microsoft Excel to arrange the clipped audio files within the frames and codes that had emerged during the course of the research, from our previous experience, and from the themes that emerged during the initial data analysis.

The structure provided by Excel spreadsheets and visual representation of the data allows us to see clearly the patterns as they emerge from the data, noting trends, areas for further study, and recurring themes. Because the actual participants' voices have been captured in the clips and have not been transcribed, we can replay their words and relive interaction with them. This tends to keep the data fresh and true, reducing the risks associated with transcription, such as misinterpretation, transcription errors, and loss of contextual cues (Merriam, 1998). We feel that it allows us to keep the richness of not only what was said but also how it was said.

In the remainder of this article, we describe and discuss three cases in which this digital audio method was used. In Case Study 1, we report on a case study of teachers who had participated in a professional development program. It was the first formal use of the analysis method suggested here, and it formed the basis of a 2004 doctoral dissertation (Childs, 2004). In Case Study 2, we report on the process used in a 2005 research study and master's thesis (Shervey, 2005). It is different from the first case study, in that we had refined the process by using QuickTime to edit the clips. In Case Study 3, we describe the process used in a 2004 research study undertaken by a large urban school board (Wheatcroft, Petrovic, & Childs, 2004).[2] It is different from the previous two studies, in that we attempted to use a Sony digital audio recorder, rather than a digital camera, to capture the clips.

Case Study 1

The purpose of this case study research was to develop an understanding of the phenomenon of an online professional development program for online teachers, called ePD, offered by Innovative Learning Services (ILS) of the Calgary Board of Education in Alberta, Canada (Childs, 2004). In an attempt to provide a holistic account of the real-life phenomenon of online professional development for online educators, we used frame and code analysis, mentioned briefly above, in this case study research design (Goffman, 1959, 1974) which was informed by principles of ethnography. This approach facilitated a "research process [that] is one of constant interaction between problem formation, data collection and data analysis . . . [and] brings a variety of techniques of inquiry into play . . . [and one in which] the observer is the primary research instrument" (Walsh, 1998, p. 221).

Childs (2004) used the case study approach to (a) remain in keeping with the philosophy of the case that was studied, the ePD program, (b) value and recognize participants' voices, (c) accommodate the

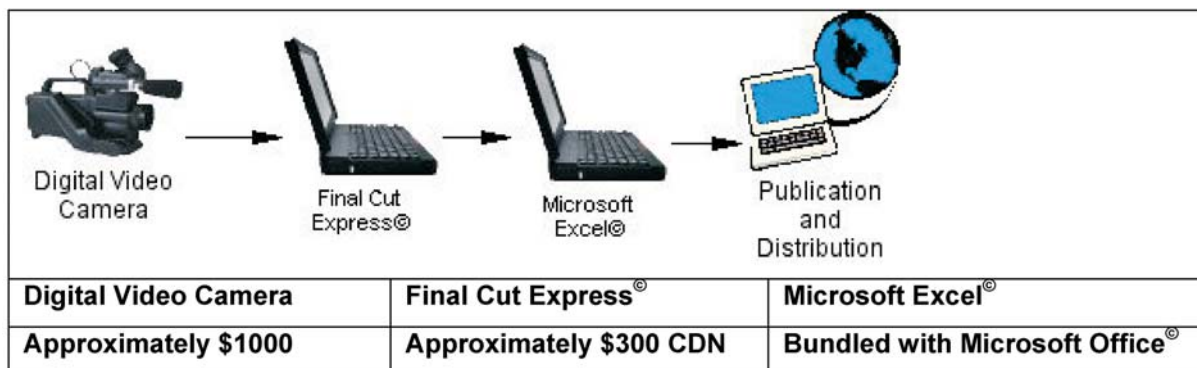| Digital Video Camera	Final Cut Express[©]	Microsoft Excel[©]
Approximately $1000	Approximately $300 CDN	Bundled with Microsoft Office[©]

Figure 1. Case Study 1 method

Theme	Research Assumptions	Based on Literature	Authors	Evidence from Participants
Capacity as online educators: online course developers and facilitators	The participation in the ePD program would result in building the capacity of its participants as online educators.	Program design and stated intentions of the ePD program.	Program developers interview; Knowles	
	Participants should be able to provide evidence of that.	Effective PD translates into practice.	Lieberman; Guskey; Fullan	
		Role and purpose of PD		

Figure 2. Sample Microsoft Excel spreadsheet analysis

cyclical nature of investigating the phenomenon of online professional development for online educators, and (d) meet the needs of ILS and the research community. The philosophy of case study research was also consistent with the constant interaction and evolving inquiry principles of ethnography, whereby "analysis of data feeds into research design; data collection and theory come to be developed out of data analysis and all subsequent data collection is guided strategically by the emergent theory" (Walsh, 1998, p. 221).

Data were collected throughout the study from ePD program documents provided by program developers, unstructured interviews, informal conversations, and semistructured interviews. Field notes were made throughout this research. Childs expanded on the field notes following interviews and informal conversations to add theoretical, methodological, and personal notes. She recorded the semistructured interviews with program developers and participants using digital video with the lens covered, as the audio was of primary concern. Two groups of people made up the ePD case: two program developers and 352 ePD program participants. These participants were eventually represented by a subset of 11 who participated in this research.

Childs analyzed data by transferring all of the digital audio data collected from the interviews into Final Cut Express, a video-editing software and database program. Final Cut Express allowed her to edit the audio files into audio clips that could be coded using the frames and codes identified from the data analysis, as not everything that was said in the interview was worth transcribing and considering. This method also allowed Childs to remove the extraneous conversation. It was originally intended that the database of Final Cut Express be used to sort, store, and house these clips. However, it was later determined that this was not within the capabilities of the software. As a result, the edited audio clips were saved as QuickTime files and exported to Microsoft Excel (Figure 1).

Childs created individual spreadsheets for each code, thereby allowing her to see the support for the frames and codes and arrange the audio clips as required (Figure 2). The clipped audio files were not transcribed until the analysis was complete and the decision was made to incorporate the excerpt into the final report. This allowed her to review and consider the actual words and intonation of the participants during analysis. As mentioned above, it was felt that this would decrease the risks associated with transcription (Merriam, 1998) and keep the richness of not only what was said but how it was said. The work of Crichton and Kinash (2003) and Kinash (2004) on virtual ethnography techniques supports this approach to data analysis, as it "allows for the construction of in-depth depictions of the events of the every day life of people" (p. 1). Once the data were analyzed and the frames and codes for this research were finalized, a selection of audio clips that gave evidence of these frames and codes were transcribed for inclusion in the dissertation.

Development of the initial frames and codes started with a review of the literature, providing "a foundation for contributing to the knowledge base" (Merriam, 1998, p. 51) and for sorting the pieces of information for analysis. The frame is defined as the "principle of organization which governs the sub-

jective meanings we assign to social events" and the code as "a device which informs and patterns all events that fall within the boundaries of its application" (Goffman, 1974, pp. 7-8, 11). As Crichton (1997) has stated, "In essence, the code is a label or description for the events within the task or activity (frame)" (p. 57). A frame then becomes a container that holds subjective meanings about that which we are seeing in the social world; the codes act as labels on the various subjective meanings within a container. The development of frames and codes is an iterative process. It is through the data analysis process, the sifting and sorting of data into themes with similar characteristics, that the initial frames and codes were established. These frames and codes could then be modified throughout the data collection and analysis processes. The final frames and codes established in this research are based on this dynamic and iterative analysis, and are supported by the interview data, the literature in the field, and the previous experience of the researcher.

As explained by Patton (1990), the "purpose of interviewing, then, is to allow us to enter into the other person's perspective" (p. 196). Because the research questions focus on the development of an understanding of how participants interpreted their ePD experience, they dictated the appropriateness of interviewing as a data collection strategy for this research (Merriam, 1998). Childs developed an interview protocol based on Merriam and followed it consistently. The prompts for the semistructured interview consisted of four questions per frame. Childs used them to explore the revised initial frames and codes as well as to create space for further discussion of the interviewees' experience in the ePD course(s). She conducted a final unstructured interview with the program developers on September 26, 2003, to discuss the status of the ePD program given the changes in the mandate of Innovative Learning Services made by the Calgary Board of Education during the research study and the expected increase in program enrollment from that mandated change.

Case Study 2

The purpose of this case study research was to explore the knowledge, skills, and attitudes of the first cadre of preservice teachers to work in distributed learning environments at the University of Calgary (Shervey, 2005). As the principal researcher was also the instructor of the course, a participant observer role was possible for this ethnographic case study.

The course ran for 13 weeks. Students worked online in both Web CT and Blackboard. They attended campus seminars, worked with partner teachers, moderated online discussions, conducted action research, and built learning objects. Because of the range of learning environments and experiences afforded these students, there were opportunities for multiple forms of data collection.

Initially, students were given a precourse survey. Their e-mail communications with the instructor were compiled, as were their online discussion postings. Two focus group interviews were conducted, and students gave a final presentation of their research and learning objects. Finally, they were given an exit survey. The focus group interviews were video recorded and the surveys were completed as text files so the content could be compiled digitally. The students submitted their research papers as e-mail attachments, and the learning objects were posted on their personal Web pages. Because all of the data were digital, the researchers could analyze them using the frame and code methods presented above and sort them within a spreadsheet.

Treatment of the audio was slightly different from that in the first case study. Based on the experience reported in Case Study 1, we decided to explore the possibility of using just QuickTime Pro rather than Final Cut Express. QuickTime Pro provided the functionality required to cut the audio files into manageable clips. However, it also required that we first import the audio files into iMOVIE to convert them into QuickTime movies, a format easily imported into QuickTime Pro. The advantages to this change include our not needing to learn Final Cut Express, the size of the audio files, and the ease of use of QuickTime Pro (Figure 3).

A further simplification of the process came with the use of an iPOD with a microphone attachment. The iPOD replaced the video camera, and participants seemed less threatened by it. It still appears that some participants view a video camera as intrusive, and they can be self-conscious while it is running, even though eventually they do seem to forget that it is there. Files exported from the iPOD can be

Digital Video Camera	QuickTime Pro©	Microsoft Excel©
Approximately $1000	Approximately $29.95 US	Bundled with Microsoft Office©

Figure 3. Case Study 2, original method

imported directly into QuickTime Pro, which saves the time of importing and exporting through iMOVIE. It also reduces the demand on hard drive space (Figure 4).

As with Case 1, the use of Microsoft Excel to manage the frame and code structure allowed us to see patterns as they emerged and to gain a sense of when the field was saturated and the data was becoming repetitive (Merriam, 1998).

In addition to ease of data analysis and management, actual clips, spoken by the participants themselves, can, as per ethical review request and if consented to by participants, be selected from the spreadsheet and repurposed for use in presentations, Web-based papers, and a range of other uses, adding an increased credibility and impact to the data that is typically presented in a transcribed format.

Case Study 3

The purpose of this case study research was to develop a collaborative, adaptive educational technology strategy for the Calgary Board of Education (Wheatcroft, Petrovic, & Childs, 2004). The Calgary Board of Education (CBE) is the second largest board in Canada, with more than 98,000 students and approximately 5,700 teachers. Two researchers from Innovative Learning Services (ILS) and the two coauthors conducted this research. We used a combination of case study and survey research to maximize participant involvement in the research.

The ILS researchers selected 17 representative case study schools or units for the research. Interviews of 1 hour in duration were conducted with the administrators at each school or unit by the two ILS researchers. In addition, one or both of the ILS researchers conducted 1-hour focus groups of teachers and/or staff members for each of the 17 schools/units. Both ILS researchers made field notes throughout the data collection process.

iPOD© and microphone	QuickTime Pro©	Microsoft Excel©
Approximately $500 CD	Approximately $29.95 US	Bundled with Microsoft Office©

Figure 4. Case Study 2, final method

All interviews and focus groups were digitally recorded using a Sony digital audio recorder, a departure from the original method of digital audio collection outlined in Case Study 1, above. This was done in the interest of cost and availability of existing equipment. All digital audio data were analyzed by the author using the method outlined in Case Study 1.

We were unable to transfer the digital audio data collected from the interviews into Final Cut Express, as the software could not import the file types created by the Sony digital audio recorder. As a result, the clipping was done using the Sony digital audio recorder software, Digital Voice Editor. This made for a more tedious process due to the lack of sophisticated audio-editing features available in this voice editing software. As a result, the clipping required us to listen to the full clip to determine where the "out" points could be set, then go back and clip the audio, in effect, listening to the clip twice at full length to clip the relevant portions for sorting and coding.

The clipped audio files were named to reflect the content of the clip. Because of the limitations of the .dvf file type, the clips could not be placed as QuickTime files into Microsoft Excel for sorting. Instead, we sorted them initially by file name on individual Microsoft Excel spreadsheets. These initial themes were confirmed by listening to the clips associated with that file using the Sony Digital Voice Editor software.

As with Case Study 1, we did not transcribe the clipped audio files until the analysis was complete. However, the process to get to this point was laborious due to the choice of: (a) the digital audio recording device, (b) the 32 1-hour clips that needed to be reviewed and sorted, and (c) the need to alternate between the Excel spreadsheets and the Digital Voice Editor software to sort through the clips. As with Case Study 1, once the data were analyzed and the frames and codes for this research were finalized, we transcribed a selection of audio clips that gave evidence of these frames and codes for inclusion in the final report and supporting Web site for the CBE Educational Technology Strategy.

Conclusion

As stated earlier, the importance of maintaining the actual voices of the participants in a case study is critical. Traditional methods of transcription run the risk of reducing the impact of the participants' words by removing them from both the context in which they were collected and the manner in which they were said. In this article, we suggest that with digital captured audio, iMOVIE, QuickTime and Microsoft Excel, the essential elements of ethnography can be supported, honoring both the process and product of thoughtful interactions between the researcher and the field.

As stated in our discussion of Case Study 2, not everything worked smoothly. The quality of the voice recorder, and its capability with editing software, is a crucial link. To date, we have not found an option better than the iPOD and QuickTime Pro combination. Although other editing software is available (e.g. Final Cut Express, Final Cut Pro, Adobe Premiere) the costs are high in terms of both initial expense and learning time.

Since this research protocol was first presented at the Sixth Advances in Qualitative Research Methods Conference in Edmonton, Canada (February 2005), a number of participants attending the session have contacted us and expressed their interest in using it. All have agreed to stay in touch and share best practices and lessons learned. A lingering frustration is the inability of Microsoft Excel to hold the icons of the QuickTime file within specific cells within a spreadsheet. However, this technical issue occurs only when the data are sorted within Microsoft Excel. There does not appear to be an easy fix for this problem.

In terms of next steps, we will explore the use of a database rather than Microsoft Excel to remedy the issue of the floating icons. In addition, we will work with our ethics boards to understand the issues associated with using audio files within conference presentations, as we recognize even more fully the power of participants' voices in the telling of their stories.

We recognize that our work is at an early stage, and we are excited by the initial findings and the fact that the methods presented in Case Study 1 withstood a doctoral oral defense and a master's thesis submission. It is currently being used in at least two other master's thesis methods. As we continue to work with this method, we see opportunities to incorporate other software applications to assist us with analysis and reporting.

We believe that if we are able to continue to work with the original source material for as long as possible, the integrity and character of the field under study will be maintained. By our doing so, those interested in the research findings will be able to hear original source material, come to their conclusions, and participate in thoughtful, critical debate of the field under study.

Notes

1. Hammersley (1990) reported on the practice of armchair anthropologists, who wrote research papers around the turn of the 20th century from the comfort of their libraries and offices drawing on the field notes of others.

2. All subsequent references to Case Studies 1, 2, and 3 relate to Childs (2004), Shervey (2005), and Wheatcroft, Petrovic, and Childs (2004), respectively.

References

Childs, E. (2004). *The impact of professional development on teaching practice: A case study*. Unpublished doctoral dissertation, University of Calgary, Calgary, Canada.

Crichton, S. (1997). *Learning environments online: A case study of actual practice*. Unpublished doctoral dissertation, University of Sydney, Australia.

Crichton, S., & Kinash, S. (2003). Virtual ethnography: Interactive interviewing online as method. *Canadian Journal of Learning and Technology*, *29*(2) 101-115.

Genzuk, M. (2004). *A synthesis of ethnographic research*. Retrieved May 30, 2004, from http://www.rcf.usc.edu/~genzuk/Ethnographic_Research.html

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. New York: Harper Colophon.

Hammersley, M. (1990). *Reading ethnographic research: A critical guide*. London: Longman.

Harris, M., & Johnson, O. (2000). *Cultural anthropology* (5th ed.). Needham Heights, MA: Allyn and Bacon.

Kellehear, A. (1997). *The unobtrusive observer*. Oxford, UK: Saint Martin's.

Kinash, S. (2004). *Blind online learners*. Unpublished doctoral dissertation, University of Calgary, Calgary, Canada.

Merriam, S. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco: Jossey-Bass.

Patton, M. Q. (1990). *Qualitative evaluation methods* (2nd ed.). Thousand Oaks, CA: Sage.

Shervey, G. (2005). *The impact of online teaching and learning on pre-service teachers*. Unpublished masters thesis, University of Calgary, Calgary, Canada.

Walsh, D. (1998). Doing ethnography. In C. Seale (Ed.), *Researching society and culture* (pp. 217-232). London: Sage.

Wheatcroft, M., Petrovic, E., & Childs, E. (2004, May). *Towards a vision of schooling for the future: Examining the Calgary Board of Education as a learning organization*. Paper presented at the BCEdOnline Conference, Vancouver: Canada.