Reflections on the Dissertation Process and the Use of Secondary Data

Rochelle Mahoe

When I was asked to write an article on how I used quantitative research methods in my doctoral dissertation, I eagerly jumped at the chance. Not because I am some sort of expert on the topic, but rather because I wish someone had given me some insights into the journey I was about to embark on. Although the course work for my doctoral degree served as a strong foundation for the voyage, the dissertation phase was still a leap of faith. I often felt as though I was "shooting in the dark" and would be lucky to hit the target. In the following paragraphs, I would like to share the personal experiences, struggles, and "aha moments" that I endured and enjoyed during the dissertation process. I hope that these thoughts will provide others with some insights and words of encouragement to persevere in completing what is undoubtedly one of the most rewarding challenges in a doctoral student's education. A good place to start, then, is to tell the story of how I became interested in the topic that would eventually become my research project.

Introduction to Problem

Developing a research problem is usually a personal process that unfolds over time. As a teacher at a local public high school, I have always been interested in how critical the student's first year of high school is to later graduation. Early in my career, like most beginning teachers, I was assigned the "lower," less rigorous, courses. I quickly became familiar with teaching freshmen, including "repeat" freshmen. This experience helped me understand that if students fell behind early in their high school years, then they would be trapped in a discouraging game of "catch-up" for the remainder of their high school career. Thus, my interest in freshmen retention began with my early experience as a teacher.

Coincidently, a few years later, my school principal shared with the faculty that for the past five years, the freshmen retention rate at our school was approximately 20 percent. Almost one out of five freshmen had not earned the required five credits to be promoted to sophomores. This regrettable figure inspired me to investigate the problem further and determine how our freshmen class consistently averaged nearly 600 students per year in contrast to the graduating senior class that averaged about 400 students. These numbers suggested that only two-thirds of the freshmen class were graduating in four years. If so, what was happening to the other 200 students? Were they dropping out, graduating late, or transferring schools? More importantly, was this problem unique to our high school, or did it plague our entire public school system? These alarming statistics strengthened my passion to learn more about this phenomenon, and it became my personal quest to find answers to the problem.

Background Information

By this time, I had completed about half of my required doctoral coursework. And because high school freshmen retention was my hot topic of interest**,** I began to incorporate the subject into my remaining course assignments. In other words, whenever it was possible, I used my courses to explore related topics to gain background information on the problem. In one of my doctoral seminars, for example, the assignment was to complete an annotated bibliography on a topic of our choice. Knowing that at some point I had to write a review of the literature in my dissertation proposal, I decided to research the topic of freshmen retention in high school. I found very little, so I widened the search to other related topics, such as social promotion and the transition to high school. In other courses, I focused attention on related areas of interest, such as school-to-school transitions and dropping out of high school. Although I was still far from developing a research question, I gradually began to assemble the kinds of knowledge that would serve me in my dissertation work. Having established a solid literature base in the preliminary stages helped me to build a strong dissertation proposal that needed few changes.

In another of my advanced research courses, I was able to conduct a preliminary analysis of freshmen retention rates using the Hawaii State Department of Education (DOE) database under the guidance of a professor who had been given access to it. Unfortunately, the numbers were not very revealing. The database showed that the greatest percentage of retained students occurred during the ninth grade year, but the percentage was too small to allow me to make any useful inferences. Perplexed by these numbers, I asked my school registrar about the large discrepancies in percentages between

the school level and the state level. She explained that some of the inconsistencies might be due to the timing of the reports. On any given day, the numbers could vary depending on when the registrars from the individual schools entered the appropriate information. Thus, the percentages varied from school to school depending on whether students were flagged as retained or recorded as promoted. In addition, the database was limited to those students who remained in the public school system until they graduated. Students who moved out of state or left for a private school were dropped from the database. This made it difficult to track whether a student dropped out or transferred schools once they exited the system.[1] This practice decreased the number of students per grade level and, subsequently, lowered the retention rates.

Need and Significance

At this stage, I needed to take a moment to reflect and reassess what I was doing. I was convinced that it was a phenomenon that was not unique to my school that high numbers of ninth graders were not being promoted to the tenth grade. In addition, I realized that because of the way that data was recorded in the DOE database, I would not really be able to use it to dig deeper into the issue and get answers to this problem. The roadblocks and challenges that I encountered, however, only served to strengthen my resolve and encourage me rather than discourage me in seeking my goal. The fact that I could not figure out what was happening to these students at the school and state level, and the lack of available literature, convinced me that this was an untapped area that needed to be explored. I believed that this was an important problem to understand in more detail and that my efforts would not be wasted. I also believed that my findings would make a useful contribution to the accumulated research on dropouts. It is important to know this in dissertation work, as it helps provide the critical element of motivation to the process. You need to feel that what you are doing is worthwhile.

My initial hypothesis was that the problem at the ninth grade was due to the transition to high school. I thought that many students probably had a difficult time adjusting to their new social and academic environment, and therefore, fell quickly behind at an early stage in their high school careers. This seemed plausible, but I also wanted to know the lasting effects of this early experience of failure. Do students ever recover and graduate on time? And if so, what are some of the factors that can be attributed to their success? These were questions to which I could not easily find answers, and they kept circulating through my mind. I questioned whether I could really find the answers. At this point, if I wanted to pursue this topic, I had only two choices: to analyze the data at the school level, or to use the data provided by the DOE. The inconsistencies in reporting data discouraged me from using the DOE's information. The alternative was to pursue a case study of my school. When I really thought about it, however, in order to accomplish my goal I would have to track one freshman class for four years. This was discouraging. In addition to the time element, I would also have to develop a survey in order to gain information that was not available from student records. I did not want to spend 5–6 years on my dissertation, so I abandoned both ideas.

These choices brought me to something of a stalemate. But I was not quite ready to abandon the topic of freshman retention. Instead, I did some additional reading on at-risk students and effective school research as I worked to expand my literature base, hoping to gain some insights into what to do next. Fortunately, a few months later, by chance, one of my professors was given information about several workshops sponsored by the National Center for Education Statistics (NCES). He noticed that one of the workshop topics included the transition to high school and forwarded the information to me. I later learned that these workshops were actually all-expense paid training sessions that showed researchers how to use the information compiled by the survey work of NCES. Luckily, there was one study, the National Education Longitudinal Study of 1988 (NELS:88) that followed a cohort of eighth graders through their transition to high school and college. Seeing that this could possibly be the answer to my prayers, I applied for the workshop in March and received my acceptance in May for a one-week training session in June.

At the training session, I once again found myself in unfamiliar territory. Many of the professors and other graduate students were already conversant with the database. And as I knew nothing about it, I spent most of the time struggling to understand why the study was conducted and learning how to access the information. Meantime, the others were more prepared to ask specific questions relevant to their purpose.

[1] Since the time of my preliminary analysis (which was over 5 years ago), the DOE has made considerable efforts to improve their record keeping in the database system. It is possible to gain more consistent information with their updated information system.

In hindsight, I should have read about the NELS:88 database on the web and requested the public access database ahead of time. If I had done so, I would have made much better use of my time and the expertise of the NCES statisticians at the training sessions. Instead, I spent about 3 months after my return trying to figure out if this database was something I could use. Fortunately, by playing with the dataset, conducting a few simple analyses and reading several articles that used the dataset, NELS:88 gradually emerged as a dataset that would address my research concerns. One of the key objectives of NELS:88 was to provide longitudinal data about critical transitions experienced by students. Although no follow-up survey was conducted during students' ninth grade year, transcripts and course information were available for every year of high school through the "restricted access" database, which I later obtained.[2] Although many studies had been conducted using the NELS:88 database; very few had been published using the restricted student files.

Once I had determined that I was going to use the NELS:88 database, I began my dissertation proposal. I expanded the scope of my literature review to also include student resiliency, positive psychology, student persistence in higher education, and small school research. By doing so, my study of a high school freshman retention problem evolved into a comprehensive examination of student persistence. More specifically, the study extended previous work by combining the psychological and sociological perspectives of dropping out and simultaneously investigating the effects of individual-level and school-level variables on students' decisions to stay in school until graduation. The multilevel study examined how school structures and processes serve as supports to students' academic and social engagement (for all four years of high school), and their subsequent influence on student persistence.

Methodology

There are many advantages and disadvantages to using secondary data. The obvious advantages in this case were the cost and time. First, the National Education Longitudinal Study of 1988 (NELS:88) conducted by the National Center of Educational Statistics (NCES) offered a source of data at no personal expense to me. This comprehensive study followed a nationally representative sample of eighth graders through their secondary schooling and post high school experiences. Millions of dollars were spent over a twelve-year period to obtain information from students, dropouts, parents, teachers, and schools to help track student achievement, educational status, and transition to school and the work place. Second, over the twelve-year period, in addition to the initial student survey given at the eighth grade, four follow-up surveys were conducted during the tenth grade, twelfth grade, two-years after high school, and eight-years after high school. In my opinion, the time it took me to learn and gain access to this comprehensive study was minimal compared with the time it would have taken me to conduct my own surveys and collect other student information.

On the other hand, there are also some disadvantages to using secondary data. One major disadvantage is that someone other than you designed the surveys. Therefore, the questions and available information are not always coded or worded exactly as you might like them to be. Although I did not have to spend time creating a research instrument, I did have to spend time fine-tuning and adjusting it to fit my needs. This is why it is critical to have a strong literature base. When using secondary data, it is essential that one should familiarize oneself with other empirical studies that have used the same dataset. This will provide information about the struggles and limitations that others have encountered. In addition, a strong theoretical base is crucial to a solid dissertation. Every decision made regarding the study should be grounded in theory. For example, the variables selected and the method in which they are recoded[3] should be justifiable. Arbitrary decisions without a theoretical basis may weaken your study or cause future headaches. The researcher should be able to give a plausible explanation for every significant and non-significant finding based on the pertinent literature.

Concluding Thoughts

Although the dissertation phase is a lonely venture, it is also the most rewarding part of the doctoral process. I would like to offer some final words of advice. First, the research will be consuming. It is critical, therefore, that you select a topic that you are passionate about and can keep you motivated. You will be your best cheerleader. Second, when you

[2] The data are available to researchers holding a license issued by the NCES. A license (control number 030227729) was awarded to Dr. Ronald H. Heck of the University of Hawai'i. In addition, this study had been determined to be exempt from a full review from the Committee of Human Subjects. In accordance with U.S. Department of Health and Human Service regulations, the University awarded a certificate of review (CHS #11529).

[3] The response for survey items are typically written as a string variable (e.g., A–F) or numeric variable (e.g., 1–5). However, based on the researcher's needs and the theoretical and empirical literature, it may be more meaningful and useful to "recode" (such as rewriting 1–5 as 5–1), "dummy" code (creating a dichotomous variable), or combine variables to create new ones.

hit a roadblock, do not get discouraged. Instead, try to understand why you are faced with this hurdle. By understanding the limitations inherent in the problem, you may learn how to get around it. Your persistence may lead you to make a future contribution to the larger body of research. Third, the role of background research is critical to any study. You need to understand the "bigger" picture before you can attack your area of interest. Look into all the related areas for insights. A strong literature base will prove helpful in the long run and offer insights that may enable you to work more efficiently. Last, use all the members of your dissertation committee. Each professor brings her or his own unique expertise to the table. Keep them informed along the way and utilize their knowledge. You will be the expert on the topic, but they will help you elevate it to a higher scholarly level.

Rochelle Mahoe received her PhD in educational administration (K–12) from the College of Education, University of Hawai'i at Mānoa in 2003. She is currently vice principal at Princess Miriam K. Likelike Elementary School on Oahu.