Harvard University

From the SelectedWorks of Anu Vedantham

Spring January, 2017

Mixed Methods Data Collection to Explore Youth Video Creation Behaviors

Dr. Anu Vedantham, Harvard University

Available at: https://works.bepress.com/anu/28/
# SAGE Research Methods Case Submission for Consideration

**Case Title:** Mixed Methods Data Collection to Explore Youth Video Creation Behaviours  

**Author Name(s):** Anu Vedantham  

**Author Biographies:**  
Anu Vedantham completed the doctorate in higher education management at the University of Pennsylvania in 2011 under the supervision of Laura Perna. This article summarizes the process of research question generation and dissertation data collection focused on gender differences in the creation of online videos by undergraduate students. Anu has directed the Weigle Information Commons at the University of Pennsylvania Libraries since 2007. She has authored several book chapters and peer-reviewed journal articles. Browse her writings at [http://works.bepress.com/anu](http://works.bepress.com/anu)  

---  

**Relevant Disciplines:** Higher Education, Sociology, Psychology  

**Academic Levels:** Postgraduate  

**Methods used:**  
- Focus Group Studies  
- Sampling  
- Survey  
- Factor analysis  
- Generalized linear models  
- Logistic regression analysis  
- Content Analysis  
- Survey Designs  

**Keywords:** social media, online videos, YouTube, Facebook, technology acceptance, self-efficacy theory, stereotype threat  


---  

**Abstract**
This case study describes the creation process for a mixed methods study from 2009 to 2011 examining for the first time gender differences in how undergraduate students create videos for online platforms such as Facebook and YouTube. I conducted this research for my doctoral dissertation at the University of Pennsylvania’s Graduate School of Education under the supervision of Dr. Laura Perna. The case describes my research design process, development of a conceptual framework, narrowing focus of research questions, collection of data, sampling and survey design, data analysis and dissemination. The case describes the process of building on personal reflections on software use and video creation, through a traditional literature review, to create a robust theoretical framework that can outlast the technologies (cell-phone cameras, video-editing tools, etc.) under study. I created a conceptual framework that builds on four theories: Self-efficacy theory, the Technology Acceptance Model (TAM), stereotype threat theory and learned helplessness theory. I collected surveys to far exceed the required sample size, and managed focus group and interviews for qualitative data to complement the logistic regression analysis. My research revealed significant gender differences in video creation that were explained by the conceptual framework. Results inform design of school projects with required video creation, and supports for video skills acquisition. Concepts include perceived ease of use, perceived usefulness and comfort with social risk.

Learning Outcomes

1. By the end of this case students should be able to appreciate how to collect data from undergraduate students.

2. By the end of this case students should be able to understand the iterative nature of research processes in higher education.

3. By the end of this case students should be understand the process of generating and refining research questions especially in the context of personal curiosity and feasibility of data collection.

4. By the end of this case students should appreciate the need to explore and master multiple software tools as part of the process of analysing research data.

Project overview and context

As part of my doctoral studies in higher education management at the University of Pennsylvania’s Graduate School of Education, I completed a research project on gender differences in the creation of videos for online consumption by undergraduate students. I conducted this research for the Ed.D. program under supervision of Laura W. Perna while working fulltime as the director of a collaborative learning space for undergraduates at Penn Libraries. I describe the process of constructing a research project that built on my personal curiosity and the resources available through my work environment to shed light on a topic relevant to new media literacies. I began with a question of personal interest, connected this question to my work and my graduate studies, created a conceptual framework, collected a large amount of quantitative and qualitative data, used
regression and coding techniques to generate research results, and after completing the dissertation, shared my findings through an open access university repository.

Early in my studies, my professors emphasized the iterative nature of education research. I initially pushed back against this sense of cyclical effort, but over time, I have found this to be a helpful mental model. My research project went through ever-narrowing iterations of reading, note-taking, abstracting, data collection, statistical analysis, consultation with experts and writing. I had to maintain flexibility throughout the process without losing sight of my overall research goals and project structure.

**Research Design**

**Building on personal curiosity**

The Ed.D. program at Penn GSE challenges doctoral students to explore research topics that begin with their own interests. Building on theories of situated learning, this process helps students with motivation and maintenance of momentum through the doctoral process. Through a series of graduate courses, students learn how to interpret personal experiences in light of educational theories about human development, organizational change and learning strategies.

As I explored my own educational experiences, I noticed ways in which my experience as a female computer scientist differed from those of my peers. I reflected on my own experiences working with new software and my observations through my work on how others interacted with new software. I perceived large gender differences in some aspects of software use, and minimal gender differences in others.

In the United States, computer programmers are 85% male. Game designers are 85% male. Women and men are just as likely to post on Facebook and share ideas and thoughts on social media. Men expressed greater levels of confidence with spreadsheet use than women. In contrast, other daily use software such as word-processing, web browsing and mobile phone apps appeared gender-neutral. I was fascinated with an initial research question that stemmed from my personal observations: “Why do men and women respond similarly to some types of software but differently to other types of software?”

I began discussing my question with classmates and faculty members in the higher education management department. They confirmed some aspects of my observations and referred me to relevant research articles. Once this question had taken root in my own mind, I began by listing observations of behaviours that are expected with different types of software.
I noted that social media tools – Facebook, Instagram, Pinterest – can be accessed quickly on your mobile phone. You can spend a minute or you can spend hours but you retain control over the time spent. Computer programming on the other hand requires long hours staring at the screen. Code needs to be debugged. Syntax needs to be mastered. User interfaces are complex, confusing and opaque. Time commitments can be unpredictable. Small tasks may take hours. You may need to ask for help often when you first begin learning this type of software. You may hesitate to even get started if the software interface looks intimidating at first glance.

Personally, I enjoy trying out a new piece of software. Poking and playing around are how I learn best. Software that is considered difficult is often more interesting to me. My own undergraduate background is in electrical engineering and computer science, and I wondered if this had given me a foundation for learning how to approach new software.

I generated a set of related questions: Why do women and men – on average – seem to approach the same type of software differently? Why do some women expect failure on first trying out a new type of software? What determines perseverance? Does practice help build confidence?

Starting with a question of personal curiosity became important to me later in maintaining momentum through the literature review and data collection processes. I approached data collection wondering what I would find, and this curiosity gave me patience through unavoidable data-cleaning tasks.

**Understanding the literature**

My advisor challenged me to identify research results to support my personal observations. My questions led to an extensive literature review that spanned several disciplines. I identified four theories that especially interested me. Self-efficacy theory (Bandura, 1997) talks about expectations of success and confidence, narrowly defined. Stereotype threat theory (Steele & Aronson, 1995) discusses the mechanisms by which our actions are affected by stereotypes for the groups we belong to.

My literature review led me back a few decades to Learned Helplessness theory (Abramson, Seligman, & Teasdale, 1978) based on research with mice that indicated that repeated failures can lead to defeatism. I found articles that applied this theory to how we may approach computer use,
by handing over control in a new situation if it reminds us of previous situations where we have failed.

The fourth set of ideas I gravitated towards was the Technology Acceptance Model (Venkatesh et al., 2003). The model looks at concepts such as perceived ease of use and perceived usefulness to explore why we are willing to spend time learning some new technologies but pass on others. It originates in the literature of business improvement, but has been applied in a few educational settings.

My four selected theories led me to new questions: Would women do better with learning complex software without others watching them navigate the learning process? Was it easier to make mistakes and play in private?

Articles on single-gender schooling and group activities were helpful at this stage. I progressed to articles on constructions of gender. I learned about non-western gender constructs, biological versus expressed gender, and other aspects of gender studies that were new to me. I explored journals in psychology and sociology, as well as journals dedicated to gender studies.

My readings led me to focus on online video creation as a uniquely new activity. Since the platforms for online video are only a few years old, students cannot benefit from the experiences of their parents and family members; they learn video-editing software largely on their own.

**Devising a conceptual framework**

The literature review led me to articles on online video creation from Hong Kong and Taiwan. I found a conceptual framework (Yang, Hsu, & Tan, 2010) that I could adapt for my study. The process of reading several dozen articles and categorizing them by the theoretical framework used, the cultural and organizational context and the technological focus was extremely helpful in narrowing my own thinking. My advisor guided me at important moments as I structured my own conceptual framework, which went through more than ten revisions. With each revision of the conceptual framework, it grew simpler in some respects and more nuanced in others. Using simple visual tools of PowerPoint and drawing software kept the framework drafts fluid. I settled on a blocked structure where variables could be brought in by stages and the framework could be evaluated through a series of experiments.
Creating the conceptual framework before making decisions on the kinds of data I could collect later was difficult for me. I struggled with the idea that I was creating a theory that would be later tested on observations, when I was not at all sure what I could observe and how I would collect data. This stage of the study was the most frustrating for me in many ways and I was unsure of my own work. Later, I was grateful for the clear-eyed support of my advisor since many aspects of the research fell into place easily once I had a solid conceptual framework in hand. Her guidance that I needed to first propose a conceptual framework based on my literature review before collecting data was crucial in the sequence for my overall research project.

**Narrowing research questions**

Following the design of the conceptual framework, I refined my three research questions, and start designing my survey for quantitative data collection. I decided to narrow the scope of my research to first-year students at the University of Pennsylvania. I used national educational databases to determine the characteristics of the entering undergraduate classes and generated a set of similar institutions along demographic characteristics. This process required collaboration with Penn’s institution research offices as well as analysis of historical admissions data. I decided to collect data at the point of entry to college, so that I could look at a narrow age-based cohort before college experiences at Penn could have affected the choices that students were making.

My research questions went through several iterations, growing narrower and more focused with each re-design, and settling at:

“1. For first-year traditional-age college students attending a highly selective research university, what gender differences exist in online video creation?

2. How do the theoretical perspectives of self-efficacy, technology acceptance, stereotype threat and learned helplessness inform understanding of such gender differences after controlling for demographic characteristics such as ethnicity, socioeconomic status, immigrant status and high school size?

3. Does the relationship between confidence using computers and online video creation vary between women and men?” (Vedantham, 2011)

Contrasting this set with my initial question emphasizes the power of iterative improvement.
Data collection and analysis

I began the data collection process by setting goals and a timeline. I found a nationally accepted survey instrument that complemented my conceptual framework and received permission from its designers to adapt it to the new field of online video creation. I drafted my survey with assistance from several faculty members and national organizations and pilot-tested it with a small group of students in summer classes. I conducted the same survey in online and paper formats so that I could ensure the reliability of the online format.

Penn has an entering first-year class of about 2,500 students. The standard assumptions of a 95% confidence level, a normal distribution, the largest response fraction possible of 50%, a 5% margin of error with this population size yielded a required sample size of 333. If I could distribute the survey to all first-year students, I would need a response rate of 13.3% to meet the sample size requirement. I knew I needed to collect all my surveys within the first month that first-year students arrived on campus so that Penn course experiences would not dilute their own perspectives on creating online videos. (Several Penn courses require online video creation as a standard assignment, a dynamic I was aware of through my work at Penn Libraries.)

I planned a concentrated burst of data collection for September 2010. I collected a random probability sample, a set of 1,000 randomly selected email addresses for the first-year students through the university registrar’s researcher services. I conducted extensive personal outreach to add convenience samples. I asked students, faculty members and administrators advice about how to reach first-year students at the point of arrival to campus. I collected ideas about incentives and using choice psychology (Iyengar, 2010) offered students the option to select gift certificates to the university bookstore, a high-end coffee shop and a campus sandwich shop.

Using snowball-sampling techniques, I networked intensively to reach first-years in different aspects of campus life. I contacted faculty who taught large first-year classes, library administrators, staff in first-year dormitories, and student leaders of clubs, fraternities and sororities. I attended a large number of new student orientation events. After receiving permission from dormitory administrators, I went personally on several evenings to first-year dormitories carrying pizza, soda and gift certificates. I used Facebook extensively, distributed paper flyers and stood in high-traffic areas around campus asking students to participate.

The mass campaign approach worked very well. Each step listed above yielded a sizable number of survey responses, and the intense burst of outreach brought in more than 821 responses. I had far exceeded the required sample size of 333 within that first month!
Having a large, stable dataset made the data analysis process simpler and more robust. I closed the online survey and began the process of coding the focus groups and interviews. I alternated between statistical analysis of the quantitative dataset and identifying themes from the qualitative dataset, and each informed the other. A substantive dataset made it possible to refine my statistical model, and ask deeper questions about specific aspects of video creation such as animation, audio-editing and video clip manipulation.

Developing the statistical models to express my conceptual framework took two months of sustained effort. I began with simple measures (mean, median, quartiles) and spent time visualizing my dataset in different ways. Slowly, I found errors that needed rectification and outliers that I needed to verify from the raw data. Eventually, I chose to create my own measure of complexity for the video-editing software titles mentioned by students. I contacted researchers in different disciplines at my university with software and function-related questions. I decided to use techniques of factor analysis, logistic regression and multinomial logistic regression. My model results were statistically significant, and the size and random characteristics of my sampling process enabled me to validate my conceptual framework at a high level.

The robust dataset allowed me to take on subtle questions. I broke down video-editing tasks into a series of sub-tasks, and I was able to look at gender differences at a fine-grained level. This revealed that creation of simple point-and-click videos was largely gender-neutral while the addition of titles and fine adjustment of audio was primarily conducted by men. These results strengthened the argument for the four theories I had integrated in my conceptual framework.

**Writing and Dissemination**

The last few months of my research project involved intense statistical analysis and writing. I refined my statistical models, and added a qualitative analysis component to the project. I defined how each video-editing software could be categorized, and created a structure for managing how video-editing is performed.

After I had successfully defended my dissertation and finalized the document, I decided to share it on the open access repository provided by Penn Libraries in addition to posting it in the ProQuest Dissertation Database. This choice helped me disseminate my results quickly. I could thank the many colleagues who helped me with data collection and analysis by sharing my dissertation with them via email. I was pleasantly surprised by how many colleagues responded to let me know they had read through my results.

**Research Practicalities**
The acknowledgements pages of my dissertation reveal the many groups of people I needed to contact at different stages in my research process. In addition to finding and contacting each expert, I also had to learn how to ask the right questions, and I found myself going back to the same experts a year later with a more developed set of questions. Building relationships with colleagues in several fields was one of the joys of this dissertation project.

In contrast, one of the challenges for me was the varied reading needed. I selected online video creation because it was of personal interest to me. However, YouTube was only a few years old when I began my research, and there were almost no studies available about online video. Instead, I had to search along a variety of related lines, and then translate the research to my context. A large number of the articles I read were not ultimately of much use for me, but I had to spend time reading and taking notes before I could determine whether they were relevant. A strong organizational process was essential for my progress. I used a simple note-taking strategy, labeling each article with a series of descriptive fields in an Excel spreadsheet and adding hyper-links to the online PDF of the full-text article. I used a detailed daily log to keep track of my progress and connect each article within a larger list of related works.

The last few months of the research project required intense focus on the data collected. I alternated between statistical analysis and writing, and once again it was an inherently iterative process. I appreciated the structured writing environment provided at "dissertation boot-camps" conducted by the Graduate Student Center, with two weeks of isolation with conscious removal of all distractions including phone and email access. Graduate students supported each other through the intense writing process, and the enforced seclusion was helpful for maintaining momentum. I participated in two weekly writing groups where one evening a week was set aside for writing in a library study room.

**Practical Lessons learned**

**New Software Tools**

My research project helped me to learn several software titles. My online survey was conducted through Qualtrics. With help from several colleagues, I created a survey that had university-approved branding, a progress bar and skip-logic so students only saw questions that applied to their situation. I customized the survey using panels of email addresses so that I could easily determine which of my colleagues’ assistance was bringing me survey responses. The survey software generates unique survey addresses and custom email solicitations, so that reminder notices would only be sent to students who had not yet completed surveys. By exploiting advanced features in Qualtrics, I was able to keep a daily tab on my progress towards the required sample size, and I could follow up promptly with my colleagues who had offered to help me publicize my survey.
Once the dataset had been collected, I used STATA to conduct the regression analyses. I used Microsoft Excel to manage both the quantitative and qualitative data. Near the end of my research project, I encountered NVivo, a powerful qualitative research software. I was not able to use it sufficiently for my dissertation, but I have had success with it for subsequent research projects and would strongly recommend it for qualitative data especially focus groups, interviews and survey data.

**Ask Everyone for Help**

When I first began the data collection stage, I was hesitant to discuss my research project with my work colleagues. I was not confident yet about what I wanted to study and why. My advisor guided me to create a simple template about my dissertation. Gradually, I created a clear “elevator pitch” about my dissertation that could be understood by a lay person. Creating this pitch took time and practice, and I was awkward and tongue-tied at first. Through enforced practice, I learned not to take tangents but to stick to a clear, simple script with no jargon or acronyms. I would begin by explaining my personal curiosity, then cover my three research questions, explain why I needed to collect several hundred surveys and then ask for specific assistance from my audience.

Sharing the rationale for my study, and my personal excitement about the project, brought strong positive results. My listeners would themselves become curious, and then offer suggestions for future research. Often my listeners would come up with creative outreach ideas.

I found it helpful to discuss my research project with just about every person I met. YouTube and Facebook are ubiquitous terms, and my listeners would quickly start sharing their favorite videos with me. After a few minutes, I could then ask them for help with reaching out to first-year students. I would end each conversation with a simple request – could you please share this web link via email with any first-year students you are in contact with?

Although many of my listeners could not immediately help me, they would often have suggestions of other people for me to contact. I used an Excel spreadsheet to keep track of my conversations, and worked systematically through all the first-year dormitories and high-enrollment first-year courses.

**Conclusions**
Reflecting on the three years spend on my dissertation research, I find it was a time of intense learning and thinking. Much of the process involved asking others for help, guidance and access to the first-year students I was studying. Key ideas I would emphasize are:

**Stay Organized**

I used a series of Excel spreadsheets to manage the process of my own research. These spreadsheets would include reading lists (to read thoroughly, to excerpt, to skim etc.). As I found a new author, I would keep track of a few articles by that author as well as who they cited. I used the Refworks citation manager to keep track of several hundred articles with tagged keywords, and as I read an article, I would use Refworks to maintain my notes online.

Later during the data analysis stage, scripts for regression modeling helped me to document the functions I selected, and the order in which I would manipulate my dataset for graphs and tables. I used Google Docs extensively to maintain writing and task logs as well as snippets of writing that could later be brought in to my dissertation document. Advanced techniques in Word were also very helpful to create a “document map” so I could quickly find sections of my large dissertation file to manage updates.

**Document Your Own Progress**

A daily log on Google Docs helped me to see the progress I was making. I would keep track of each person I contacted, each article I read and each hour I spent on my dissertation. Each daily entry would start with my plans for that day, and would end with my plans for the next day. The accumulation of effort would help me maintain my interest in my research; I could see clearly how my “baby steps” were moving me a bit forward each day.

**Stay Flexible**

Leaving yourself open to new ideas throughout the research project is difficult, but rewarding. My conceptual framework worked quite well for much of the study. At the very end of the study I encountered some surprises. I had to go back to extending my literature review to learn about expressed versus biological gender, and theories of gendered behavior that were not part of my four theories. I had to adjust the quantitative results for new ideas that came from the focus groups – women perceived leisure time and play very differently from the men. Playing with video-editing software was a theme that was not part of the conceptual framework and it emerged from the student interviews. This led to a substantial discussion in the next steps for research section.
Finding new ideas at the very end of my study was difficult because it raised questions about my conceptual framework. I had to remind myself that this was a small, constrained study, and I could explain the limitations of my conceptual framework as part of the concluding chapter of the dissertation.

**Share your results through Open Access**

I strongly recommend sharing your research results using the Open Access repository at your university’s library. With librarian assistance, my dissertation is now available for all to access at [http://works.bepress.com/anu/11/](http://works.bepress.com/anu/11/) and also available through Proquest’s Dissertation Database. Open access repositories can help you share your research quickly, and can help increase the visibility of your work to policy makers and educators. I have been surprised by the many hundreds of downloads of my dissertation each year, and the followup requests for speaking engagements including a paper I presented at the 2014 AERA annual conference (Vedantham, 2014).
Exercise and Discussion Questions:

1. What did you find notable about how the researcher’s questions evolved over time? Do you perceive similar changes in your research questions?

2. How did the researcher maintain momentum of data collection? Would you expect similar challenges in your data collection?

3. How important was the development of a conceptual framework for this study? How does a conceptual framework help define the research steps?

4. The researcher emphasizes the iterative nature of research. Has your research followed an iterative track? If so, how do you maintain your motivation and organization through the cycles of iteration?

List of Further Readings:


Links to any relevant web resources:

Bibliography of references cited in text:
Bibliography


