DON’T LOOK A GIFT HORSE IN THE MOUTH: WHEN THE DATA YOU RECEIVE IS NOT THE DATA YOU WANT

SAVANNAH L. KELLY

INTRODUCTION

What do you do when the intervention you set out to measure did not make a difference, and the intervention you did not want to measure was more significant than you expected? As instruction librarians, like other teaching faculty, it is critical that we understand our contribution in the classroom: What is effective, and what is not? But unlike teaching faculty who have a full semester with their students, instruction librarians often work within the narrow constraints of a fifty-to-ninety minute one-shot session. It is therefore all the more pressing that we understand best practices to evaluate both our pedagogy and students’ learning experiences. One helpful way to assess our impact is to employ a quasi-experimental research design. Experiments in the classroom are most often used to measure the effect of a particular intervention. In this case, the intervention was an online video series that students watched in addition to receiving an in-person library session. What happens, though, when you advertise and execute a full-blown research study and the intervention measured was completely negligible? What do you do then? This paper will share the author’s experience of executing a quasi-experimental design with 400+ first-year writing students and the dilemma of unwelcomed research results. Roadblocks, lessons learned, and takeaways for other instruction librarians are provided.

THE INITIAL IDEA

This project was borne from the idealism of the IL Framework and the realism of the one-shot classroom. One of the Research as Inquiry dispositions demarcates that learners “value persistence, adaptability, and flexibility and recognize that ambiguity can benefit the research process.” Likewise, the Searching as Strategic Exploration dispositions encourage learners to "exhibit mental flexibility and creativity", “persist in the face of search challenges” and “understand that first attempts at searching do not always produce adequate results” (Association of College and Research Libraries, 2016). These dispositions point to the broader role of affect in the classroom. Information literacy is not simply about doing something but about thinking and feeling a particular way. Even novice librarians will pick up on the role of affect in students’ willingness (or not) to engage in information literacy instruction. Teaching students is more than demonstrating a technique or skill set; it also involves changing students’ mindsets and their emotions.

But how is this accomplished during a one-shot session? How can librarians help students draft research questions, develop keywords, search databases, evaluate content, incorporate sources and address students’ affective processes during a fifty-minute session, especially when twenty of those minutes are dedicated to active learning? How is it possible to demonstrate the inherent messiness of research (e.g., confronting dead ends and missteps) when doing so in the one shot might confuse or discourage students? This line of questioning led me to consider alternate ways to talk directly to students. Rather than trying to compress everything into an in-person session, was there another way to provide supplemental instruction? As most teaching librarians know, the answer to that question is yes. And one of the ways you can offer additional content is through online multimedia. At the University of Mississippi, despite our enrollment of over 24,000 students, we have been slow to embrace online video tutorials in lieu of face-to-face one-shot instruction. Given that most Research & Instruction Librarians at our institution average 100 sessions annually, it was time we considered increasing the role of video tutorials to offset our unsustainable approach. Rather than immediately replacing the one shot with online content, we wanted to pilot a video series as supplemental instruction to the one shot.
I decided to create videos addressing some of the most common research struggles undergraduates experience as first-year students. Many of these issues related to students’ cognitive and emotional approaches. Returning to the IL Framework, I considered whether students “persist[ed] in the face of challenges” or “understood that first attempts at searching do not always produce adequate results” or even demonstrated “mental flexibility” when confronted with contradictory evidence. No, often students did not, and I wanted the videos to encourage students to persist in the face of difficulties. I reviewed articles and online recommendations for creating tutorials, wrote the scripts, and recorded screenshots with voiceovers using ScreenFlow. Then I created a fourth video using a green screen in the library’s media studio to personally introduce the video series. It would take viewers approximately seven minutes to watch the entire series.

1. Introduction to Video Series, 0:35
2. The Struggle with Searching, 2:36
3. The Struggle with Evaluating, 2:15
4. The Struggle to Care, 1:11

Now that the videos were created, I had to assess their impact as supplemental instruction to the one shot. The purpose of the videos was to encourage students to think and feel differently about the research process, so the measurement tool had to reflect that type of change as well. My approach was to create a research self-confidence scale based on Bandura’s self-efficacy scale guidelines (Bandura, 2006). The scale prompted students to rank (0-100) ten declarative statements related to common research challenges. For example, it asked how confident a student felt that he/she could “Get myself to continue looking for relevant sources when I can’t seem to find what I need.” I piloted the scale with lower-division library student workers before finalizing.

I hypothesized that there would be a significant difference (p < .05) in self-confidence scores between students who watched the supplemental videos and those that did not. This would be measured through the scale, which students would take twice during the semester. When using a quasi-experimental research design, it is essential to remember two concepts: measurement and control. You want to measure your constructs appropriately and you must control any other influences that could affect your outcome. It was important to take the scale twice to establish students’ initial baseline. Without that pre-scale measure, it is difficult to understand students’ gains in the post-scale measure. Likewise, it was also critical to establish a control group who would not be exposed to the intervention (e.g., the video series). Having a control group strongly increases the internal validity of the quasi-experimental study. If both the intervention and control groups are treated the same, except for the exposure of the intervention, then one can be more confident that the intervention was responsible for any changes in the outcome.

True experiments rely on randomization of individuals to groups. Since librarians teach with pre-established intact classrooms, such designs are considered quasi-experiments, which although not as rigorous as a true experiment, still enables us to understand whether or not a particular intervention made a difference between groups.

THE STUDY EXECUTION

The study was designed with first-year students in mind, but the sample had to be large enough to confidently run statistical significant tests. Fortunately I was the designated librarian for FASTrack, a Learning Community of 400+ students at UM. These students would enroll in 22 identical sections of First-Year Writing I (WRIT 101) in Fall 2016. I thought this cohort would be ideal since students would essentially have no prior library instruction and I could introduce their first exposure to the library, thus naturally controlling intervening variables. Once I identified my target population, I discussed the study with the Department of Writing and Rhetoric’s FASTrack instructors, and with their approval and buy-in, submitted an application to our Institutional Review Board (IRB) to proceed.

I randomly assigned each class to the control or intervention group. There were ten classes who served as the control, and twelve classes who served as the intervention group. Since the WRIT 101 participants had historically valued their fifty-minute instruction session, this continued to be an essential component of the study. I created an engaging curriculum for the one shot that I would teach 22 times, once to each of the FASTrack sections. In order to maintain consistency, I was the only librarian to work with these classes.

Another way I sought to control the environment was access to the video series. Those in the intervention group watched the video series once during regular class time with their writing instructor. This arrangement was to ensure that the control groups would not accidentally become exposed to the videos otherwise. The final way I controlled the research study was through a regimented timeline I provided to the writing faculty. I asked the WRIT 101 instructors to administer the research self-confidence scale to all students, both control and intervention groups, during the class prior to their scheduled one-shot session. In addition to the administration of the scale, on that same day, those in the intervention group would also watch the video series after responding to the scale. The final step was for the writing instructors to administer the scale to all students a second time. This happened on the day students turned in their source-based writing assignment.
The Results

Once I input all the pre- and post-scale data in SPSS, I realized I was not able to compute the statistical test (paired t-test) I wanted because I had not matched the pre- and post-scales (students responded anonymously). Although it was not ideal, I was still able to run an independent t-test comparing the control and intervention groups’ post-scale scores. This, however, was discouraging, as no difference existed in self-confidence scores between students who watched the videos and those that did not, which is what I had set out to assess in the first place. Out of curiosity, I decided to rearrange the data in order to run an independent t-test comparing all the pre-scales to all the post-scales, regardless of group. This showed a significant difference: students’ scores sharply increased from the pre- to post-scale measurement. This suggested that there was an alternate intervention, likely the one-shot instruction session, which had influenced students’ research self-confidence levels. This finding was problematic for two reasons. First, there was no control group to compare with since all 22 sections participated in a one shot. Thus, I could speculate that the one-shot session made the difference, but I could not be sure. The second, and more disconcerting issue, was that instruction librarians in my department were already overrun with face-to-face requests and we hoped to transition to online instructional content.

What Did It All Mean?

I was so confident that the videos would make a difference that the outcome of the study forced me to question everything. The writing faculty had followed the protocol exactly, so I immediately felt that I had designed the study poorly. The first consideration was the videos themselves. Were they too advanced? I was trying to communicate to students how to overcome common research struggles in the series, yet these students had not yet started working on their source-based assignment, nor had they even been exposed to the one-shot instruction session. They had limited contextual understanding. It was an obvious cart-before-the-horse problem.

A related issue was students’ extremely high confidence levels (90 – 100 scores out of 100) in the pre-scale. Again this overconfidence was possibly a reflection of students’ inexperience with the research process. One of the writing instructors provided an apt analogy: ask non-parents to rate their performance as potential parents (I’d be great!), and compare that to actual parents (My parenting is ok most of the time). It probably would have been wiser for me to pilot these videos with students who had at least some experience confronting research challenges.

I was also concerned that I did not measure the outcome (dependent) variable well. Was the research self-confidence scale internally valid (measured what it supposed to)? I didn’t know that for sure. Perhaps I should have spent more time developing the scale. Also, the students were particularly engaged with the curriculum during the one shot. Did the videos have a priming effect for the face-to-face session? Would it have made a difference if the one shot was boring? Finally I wondered what difference, if any, there would have been had I used the paired t-test in SPSS. Again, I just didn’t know. There were so many unknowns at the end of this study that I felt the results were useful but not definitive. I learned much more about the difficulties of designing a quasi-experiment than I did about the usefulness of the videos.

Takeaways

One of the biggest takeaways from this study was my reflection on the differences between measuring one’s impact in the classroom using basic assessment techniques (e.g., one-minute paper) and the demands of executing a rigorous social science method like the quasi-experiment. I spent so much time trying to isolate my variable from intervening factors that I wondered if I had ultimately created an artificial environment. Students did not have the ability to pause (or slow down, or speed up) the videos, nor could they watch the tutorials more than once. Did that reflect how students used multimedia in “real life”—at a set pace, only once, in the classroom? Not likely. Should I have been surprised that the videos did not impact students’ research self-confidence? Given the way I designed the study, probably not. What it does suggest, however, is that we are called to be thoughtful in our assessment approach. When you recognize flaws in your design—whether in informal or formal assessment methods—you should carefully consider your results. In my case, I will acknowledge that the one-shot most likely had a positive impact on students’ research self-confidence, even without the presence of a comparison control group because the increase in scores was so significant. Given this interpretation, the one shot will remain an integral piece of our instruction program, as much as we are able to support it. On the other hand, with respect to the flaws inherent in this current study, I am not ready to accept that video tutorials are not powerful enough to impact students’ research self-confidence. It is possible to embrace both the rigor of a quasi-experiment and the realities of how students’ engage with video tutorials. What it takes is a thoughtful and creative approach, and next time, that is what I hope to do.

References