

Rhetorical Grounding and an Agile Attitude: Complex Systems, Multi-Genre Testing, and Service Learning in UX

Laura Cunningham

Administrative Law Judge
Montana Dept. of Revenue
2 Pinecone Draw
Clancy, Montana, USA
l.cunningham@ttu.edu

Russell Kirksey

Lecturer
Department of English
Texas Tech University
PO Box 43091
Lubbock, Texas, USA
russell.kirksey@ttu.edu

Ashlynn Reynolds-Dyk

PhD Candidate
Texas Tech University
PO Box 43091
Lubbock, Texas, USA
ashlynn.reynolds@ttu.edu

Nancy Small

Instructional Associate
Professor of English
Texas A&M at Qatar
PO Box 23874
Doha, Qatar
nancy.small@qatar.tamu.edu

Charity Tran

PhD Student
Texas Tech University
PO Box 43091
Lubbock, Texas, USA
charity.tran@ttu.edu

Virginia Tucker

Director of IDS Programs
Old Dominion University
Norfolk, Virginia, USA
vmtucker@odu.edu

Abstract

This article describes a graduate-level service learning usability project and illustrates the practical tensions generated by increasing complexity in technology design. Although a foundational body of usability research prescribes standards for high-quality test design, such guidelines may not be applicable or practical in a particular context. Usability students and professionals benefit from allowing themselves the space for adaptation in order to conduct rigorous testing and meet client needs. An agile attitude—theory based but not theory bound—emphasizes user-centeredness, functionality, customer collaboration, and flexibility. Such an attitude emerges from basic principles of rhetorical theory, such as timing (*kairos*), user-centeredness (audience), genre, and rhetorical context. Bridging industry and academia, service-learning projects provide a compelling opportunity to experience the tensions and decision making required at the nexus of theory and practice.

Keywords

usability, theory, service learning, rhetoric, agile design, rigor, think aloud protocol, mental model mismatch, genre



Introduction

As PhD students in a hybrid online/residence program, we were assigned to be a usability testing team in a service-learning project where we experienced the excitement and challenges of rapid-prototyping and small-shop product development. The product inventor/manager and designers were members of our academic unit, and one of the investors was our instructor. Some of our student and faculty colleagues were considered potential real-life customers. Still and Albers (2010) assert that, “caught so often in the middle between products and their consumers, regardless of the environment, we find it almost an essential part of our job to always be ready to embrace and operationalize new methods or theories” (p. 190). Indeed, as student usability testers, we were “caught” among the product developers and investors, potential clients, graduate faculty, and a condensed timeline. We were also caught in a methodological bind because our contexts did not allow for the smooth application of theories and practices as explained in the UX scholarly literature. Our usability testing experience demanded quick, sometimes improvisational work even as we struggled to establish our footing on usability studies ground. As an outcome of that experience, this article reports and reflects on two issues: (a) how usability researchers both respect and subvert a discipline’s “best practices” based on the rhetorical nature of usability testing and (b) how embedded service learning can be a strong academic and industry bridge connecting the ideals of theory and best practices to the practical realities of serving a client.

Research Context

The six of us met as we attended our PhD program’s annual residential requirement, a compressed two-week semester during which we completed a three-credit course in usability studies. Although one of us worked with clients on website development, none had previous coursework in usability testing. We met at least four hours per day, six days per week in addition to other in-residence requirements such as attending scholarly presentations and professional development sessions. Time constraints meant the usability formal presentation and report were due to the client at the last meeting of the abbreviated timeline.

In the service learning component of our course, a start-up company, owned by a faculty member in our graduate program and staffed by two other graduate students (one in our program and one in computer science) asked our team to perform a usability test on its new product—an eye-tracking product designed for usability testing practitioners. The product is a tri-part, complex system: a camera and light device worn on the head, a capture program to record eye movements of participants, and analysis software that reveals eye-movement patterns in the screen captures. Conceived as a lower-cost competitor in a market dominated by high-end (\$20–30K) systems, the product was on a fast-track schedule to be unveiled at a conference and to go into retail production in less than a month, and our team was expected to complete usability tests and give recommendations that the clients could act on immediately. In fact, as we were testing the prototype provided on the first day of the project, a second-version prototype was concurrently being developed. The designers often responded to our informal, in-process testing results overnight.

The complex nature of the product called for a complicated set of tasks, including concurrent hardware and software testing (headgear and user interface), an approach we call “multi-genre” testing. From the initial meeting with the client until our final recommendations, the situation demanded that we extend theoretical spaces in usability testing in order to balance the needs of the client with unique circumstances and a short time frame. Our test design—including tasks, scenario, and script—required us to understand and acknowledge tensions in the UX field. As each contextual challenge arose, we sought practical but justifiable solutions pushing theoretical boundaries.

In this article, we explore the complex nature of our project on theoretical and practical grounds, focusing in part on the rhetorical features of the situation. We then describe the effects such complexity had on the design of our usability testing procedure, the outcomes of the test, and the lessons we learned regarding agility and the value of a service project bridging the academic and industry divide.

Rhetorical Theory as a Standpoint for Usability

As Redish (2010) noted: “Understanding complexity—how to create products and processes for complex situations as well as how to build usability in and assess the usability of complex systems—is becoming a major topic for technical communicators and usability professionals” (p. 198). Although we did not realize it at the time, because we were focused on meeting both graduate instructor and client demands, our academic grounding in rhetoric and technical communication served as an excellent nexus to cope with a sprinter’s deadline, a complex system, and a constantly shifting path. The rhetorical situation, defined by Bitzer (1968) as a “natural context of persons, events, objects, relations, and an exigency which strongly invites utterance” (p. 5), provided the assumptions and constraints under which our team worked.

The “natural context” included the convergence of client, designers, and testers during the two-week seminar. The service-learning project produced our exigency, and the client’s authentic need for user experience feedback provided the kairotic moment. Based on an expected level of user experience, the graphical user interface (GUI) was built with no help files or user’s manual. Designers felt the GUI was intuitive, expecting the product’s target audience—usability practitioners—to have pre-established mental models based on a popular usability-related software package already in use. Ideally, to test whether or not users had semantic structures to intuitively use the GUI, we would have given participants a Bitzer-like exigency in the form of a realistic task assignment. In other words, if our target users were UX testing professionals, then they could have already known how to competently use the prototype under development. Instead, because the participants were to interact with the prototype for the first time, we created a scenario, script, and tasks that were more exploratory and open-ended, allowing participants to create their own meaning—a rhetorical situation more aligned with Vatz’s (1973) idea that rhetorical situations are entirely subjective, that each rhetor creates the exigency (rather than wait for a natural occurrence) and then crafts a response to it. Adopting this approach to usability testing, we found the spaces between the objectivity of Bitzer and the subjectivity of Vatz. Our foundational positioning supports Yates and Orlikowski’s (2002) view “that researchers should turn their attention to the active shaping of kairotic moments” (Artemeva, 2005, p. 394) rather than passively accept the natural situation predicted by the client.

Another element of rhetorical theory is genre. When an “utterance” or response is made, it must come in some form or another. More formally, genre is defined as “typified symbolic actions in response to recognizable situation types... [that] constitute human activity” (Artemeva, 2005, p. 392). “[G]enres do not exist independently... [but rather] must be understood in relation to other genres and their interplay” (Artemeva, 2005, p. 393; see also Devitt, 2000). The way that an actor responds to a rhetorical situation is shaped by the forms of communication already in play. Complicating the rhetorical situation for our testing, the product contained three components, each one in a different stage of development (see Table 1). The hardware was a late stage prototype, and the GUI was in what Still and Morris (2010) described as a medium-level phase. The GUI had a basic look-and-feel already established and was coded for basic functionality, but it included no user support (no quick-start guide, no wizard, no built-in help). According to Lepore’s (2010) design continuum, the system was inclined towards the “prototype” level but still in the refinement or “dress rehearsal” phase. It included basic structure and flow but lacked details that would allow formal prototype testing. User documentation (a quick-start guide, help menu, embedded FAQs, or wizards) was not present. Inherent tensions were created through attempting to test the hardware, software, and documentation simultaneously. Despite being at different stages of the development process, the system could not be divided and independently evaluated as stand-alone components.

Table 1. Product Status

System component	Status
Hardware/headgear	Late stage prototype
Software/screen capture & analysis tools	Medium-level phase
Documentation	Not provided

As a further complication, the product was designed to have two simultaneous audiences or users: a usability lab professional and a testing participant. The usability lab professional would fit the headgear/camera over the participant's eye, calibrate the hardware to the software using the GUI, and start the eye-tracking process. However, because the product was designed to fit the functionality and pricing structure needs of a "small shop" usability lab, we could not presume a two-user situation. As Redish (2010) argued, "Complexity is not so much an attribute of a product or process itself as it is an attribute of the interaction between that product or process and its users. Thus, complexity is audience specific" (p. 199). Although the product system did not necessarily meet all of Redish's (2007) list of complex system characteristics, the hardware and GUI components at different stages of development and the compressed usability testing period created overall system complexity. We likened these different stages of development (and the correlation of development stages to different standards of usability testing) to a challenge of having multiple genres but only one rhetorical situation.

Thus, when presented with the rhetorical situation, including the testing variables and the client's needs, we could not rely on a single genre of usability testing. UX scholarship has generally either focused on the genre of early prototype testing or the genre of late prototype testing. The situation (exigency) driving our project was the faculty/client's need for immediate feedback, and we were not in a position to ask the designers to stop their work on the hardware and software in order to produce documentation. As students, we did not have the agency to align the levels of development of the various components. Therefore, we turned to our shared background in rhetoric and critical thinking in order to adapt the testing design to the complex product.

The Rhetorical Situation and Think-Aloud Protocol

One primary effect of our testing's rhetorical situation was a struggle with the design of our think-aloud protocol. As Nielsen (1993) noted, "Thinking aloud may be the single most valuable usability engineering method" (p. 195), and concurrent think-aloud protocol (CTA) methods have generated significant discussion regarding rigor and best practices, particularly regarding tester-participant interaction. Ericsson and Simon (1984) originally argued that rigor is primarily achieved when the tester fades into the background and resists intrusion. Nielsen (1993) was pragmatic, allowing tester interference in cases where the user was stuck or was encountering problematic areas consistently encountered in previous tests. Tamler (1998) suggested designing a test's CTA based on testing goals and recommended probing questions, active listening, targeted user guidance, and user design input as potentially useful in qualitative studies. An alternative approach grew out of speech communication theory and the notion that "the issue in usability testing becomes not one of disappearing from participants, but rather one of creating a highly asymmetrical speaker/listener relationship" (Boren & Ramey, 2000, p. 267), placing the product in the "subject" position and the participant in the "expert and primary speaker" position. Makri, Blandford, and Cox (2011) reflected on a variety of think-aloud approaches, highlighting decisions concerning when the thinking occurs, either during or after the tasks. Although emphasizing the need for authenticity in their testing's rhetorical situation by making the test sessions "as true to life as possible (within our study's constraints)" (p.341),

Makri, Blandford and Cox did prompt for data, thereby aligning with Boren and Ramey (2000) as well as with Tamler (1998) on the debate concerning researcher interference.

Barnum (2011) emphasized the discovery process—learning “what works for your users, as well as what doesn’t work, what pleases, what puzzles, and what frustrates them” (p. 10)—and affirmed a modified CTA method, based on Boren and Ramey as a field standard (see Barnum, p. 207–16). She goes further in offering advice on interacting with the test participant by adapting the Society for Technical Communication’s “toolkit” (n.d.) procedures for successful think aloud protocol: offering neutral prompting (based on tasks not features), echoing, conversational disequilibrium, and summarization at key points as “tried-and-true” techniques for think-aloud facilitation (Barnum, 2011, p. 212–214). These options certainly go beyond “um hmm” utterance limits of more conservative CTA methods but also stop short of allowing the facilitator to ask probing questions about expectations and product design advice in the face of user frustration. Considering our rhetorical situation, including the multi-genre product and mental model assumptions (discussed in the following sections), our team agreed that being able to ask questions yielded much richer testing outcomes. As Sasaki (2008) argued, think aloud is a socially-situated construct, and all of our participants tended to turn physically towards the facilitator after reaching a point of frustration with the product. Because of our particular context—limited time, a trained but familiar participant pool, a product still under development—our decision in favor of carefully targeted and worded participant-facilitator interactions ended up proving helpful because we needed reflection to understand where the task failure and user frustration originated.

Rhetorically Situated Test Design

The product designers provided the team with two key factors affecting product design, both of which influenced the product’s rhetorical context and, therefore, the study’s design: (a) the menus were structured to match those of TechSmith’s *Morae* in order to enhance ease of use for usability experts familiar with that software and mental model, and (b) the product’s price point was intended to increase the accessibility to eye-tracking systems for new or small usability labs, as well as individuals interested in eye-tracking for other fields. Considering these two assumptions together revealed a likely problem: new, small-shop, or extra-disciplinary users may not have a familiarity with the *Morae* mental model. Nielsen (2010) described such a gap in designers’ and users’ mental models, where the system may have a wonderfully rich model that is inefficient for users who each apply a different mental model when navigating that system and predicting what selections to make in the completion of a task. This was one of our concerns from early in the process and influenced our participant screening.

Scholars have debated the appropriate number of participants in usability tests (Caulton, 2001; Faulkner, 2003; Nielsen, 2000; Spool & Schroeder, 2001). Caulton (2001) described two types of variables affecting the number of users tested: user group and problem type. The product users have a common, clearly-defined background with a certain level of computer literacy, so our research sample could be homogeneous. In addition, because we sought a broad-based understanding of the product, we were looking for shared (not unique) problems. For an accurate representation of the target user group, we recruited six testing participants from our graduate program: two faculty members, two doctoral students, and two master’s degree students. Academic research institutions are one of the product’s target markets, helping to support the validity of finding participants in this pool. We also ensured that all participants had no prior experience with any iterations of either the hardware or software being tested. Participants volunteered their time and were not compensated. The department’s usability lab was covered under a blanket Institutional Review Board/Human Subjects Review approval.

Although homogenous in broad disciplinary interests, they had varying levels of usability experience. Participants averaged a 3.67 on a 5-point scale in self-reported usability software expertise (where 5 is *extensive*), with at least two participants having had direct experience with the *Morae* software. In contrast, participants reported very little experience with eye tracking, an average of 1.5 on the same 5-point scale. Without experience in eye tracking, the expectation was that users would apply their pre-existing familiarity with other usability technologies (i.e., mental models) in order to navigate the product. Formal mental model matching approach can be too time-consuming for usability study (see Sutcliffe, Ryan, Doubleday, & Springett, 2000), so we employed a more concise method where participants

were asked to freely explore the GUI for 3 minutes as they talked aloud, performed a series of tasks, then reported menu expectations and recalled navigation in a retrospective questionnaire (see Table 2).

Table 2. Testing Procedure Overview

Testing phase	Time allotted
1. Welcome, non-disclosure, consent forms	5 minutes
2. Pre-test survey on usability and eye-tracking experience	5 minutes
3. Product testing through think-aloud protocol Video recorded and coded using Morae	30 minutes
4. Post-testing survey and reflection	15 minutes
5. Exit interview	10 minutes

Although we wanted our testing to be as natural to the rhetorical situation as possible, the product's levels of development were limited to the scenarios or tasks we could test. Our scenario took advantage of the participants' familiarity with usability testing but less common knowledge of eye-tracking systems. Scripted and delivered by the same facilitator to all six participants, it began with this introduction:

This is your usability lab. You've heard of eye-tracking products in general before but have no experience using them. You've done some research and decided to purchase this product. A lab assistant has already loaded the software onto this computer and has unboxed the headgear for you. You've come into the lab to look over the product and to give the product a test run.

Participants chose their own online activity to eye-track once the device was set up, something they would have tried in a realistic setting. The online activity was discussed at the beginning of the session, in response to the question, "If you were going to complete a simple test-drive task to see how the eye tracking worked, what would you do?" The typical participant went to a favorite website and looked around for up to one minute. After this preparation, the six tasks in the script were directive (Table 3).

Table 3. Script Tasks

Task	Instruction
1	Explore menus in the interface and talk about your understanding and impressions.
2	Configure the software for this trial run by creating a project and task.
3	Try on the hardware and talk about your initial impressions.
4	Using the tips sheet provided, align your eye to the camera using the interface.
5	Start a test and follow the calibration instructions.
6	Complete the online activity you described and end the test.

Because these tasks use directive wording, we acknowledge that we were partially testing the clarity of the task instructions. Without any drafted documentation (and because of the potential for mental-model mismatch), telling participants to complete high-level tasks without providing sub-steps was not feasible. The complexity of the system meant that providing limited direction to participants allowed the facilitator to act as a verbal “quick start guide” and focused our observations on how easily the participants were able to complete the tasks.

While these interventions conflicted with standard best practices (Barnum, 2011; Boren & Ramey, 2000), the subsequent observations seemed to generate potentially helpful feedback to the GUI and hardware designers as well as to the future documentation writers (see the Results section). Interruptions were kept to a minimum and were carefully framed to reduce bias and maintain asymmetry. For example, during task one, the facilitator asked, “Do you find commands where you expect them?” and “Do you find the menu arrangement efficient and effective for your expectations?” During task three, the participant was prompted to “tell me how the headgear feels on your head.” Other than these planned, specific questions and the token verbal cues (e.g., “mm hm”), participants were occasionally reminded that they were the experts and the product was what was being tested (as recommended in Boren & Ramey, 2000, and in Barnum, 2011).

To combat bias and to address potential questions about the reliability of our data, our team addressed a number of Creswell’s (1998) procedures and all of Hughes’ (1999) areas. First, to ensure that our test reflected actual user activities, we made the scenario as natural to the anticipated user environment as possible by taking into account the product, audience, and what that audience would do when first interacting with the product. We let the participants follow the tasks with little interruption (as per the scenario) but then sought out rich feedback when a problem was encountered. We used Morae to capture video, audio, and screen movements during the tests, and three team members encoded and cross-checked their observations. The in-room observer provided an additional perspective on body language and interaction, and the combination of recordings and observations allowed for prolonged engagement and persistent observation, triangulation, and rich description. Next, because the product tested a usability lab product, testing it in a university’s usability lab provided an inherently natural environment, which suggests transferability (Hughes, 1999). Finally, with regard to dependability, Hughes recommended safeguarding against biases by having a team formed from different functional perspectives. Although at the time, our team was all part of the same PhD program, we came from various backgrounds: three of us worked in post-secondary instruction and administration, one in both secondary and post-secondary instruction, one in state government as a technical writer, and one in digital strategy and technology development for the civic sector. The mix of expertise, as well as triangulation among our data sources, enriched our analysis as we sought to further enhance the study’s rigor.

UX Testing Results

Usability testing outcomes produced hardware, software, and calibration data that we presented to the client. Our recommendations included suggesting that the client increase detailed documentation for users in all three areas.

Hardware

Participant experience with hardware was mixed. Five out of six participants thought the headset was comfortable and that the camera affixed to the headset was not obtrusive. However, concerns were expressed regarding the headset’s stability (5/6) and insecurity with positioning of the hardware (4/6). An unexpected question revealed by participants was whether or not the product would work with eye problems, including corrective lenses (3/6) and cataracts (1/6). While participants understood what it meant to align the camera to their eye, the process itself did not work as designed. Participants cited problems adjusting the headset (4/6) and physical fatigue (2/6). Ultimately five out of six participants were unable to accomplish the alignment task in the allocated five minutes and required the help of a technician in order to continue exploring the product system.

Software

Menu architecture was judged by the users as familiar and, at first glance, intuitive. Based on their think-aloud commentary and later feedback, the initial comfort resulted from the menus

being based on a common standard Microsoft Windows architecture. However, the actual GUI use and attempted completion of testing tasks created uncertainty about menu order and hierarchy. Participants noted confusion about the sequence of steps. One noted "I would think tasks would be subordinate to file, the container for the project, and hence tasks would be under file instead of separate." Although our participants had experience with usability testing, and some had direct experience with the Morae system model, they expressed unfamiliarity with the Morae-based terms: "projects," "tasks," and "tests." The presence of a dark and pixelated capture screen on the computer further aggravated the situation. In short, a combination of reliance on the Microsoft Windows mental model and targeted users' lack of expected familiarity with Morae terminology contributed to a high degree of task failure and user frustration. A lack of documentation or wizard guidance was cited as a key issue.

Calibration

Perhaps the most frustrating aspect to participants was the calibration process required to establish a connection between the hardware and the software. While they reported clearly understanding the process, participants were confused about calibration and expressed negative reactions. The software offered a "Calibration Score" which showed "0" for all participants throughout multiple attempts. Participants inferred that this response was a personal failure ("It seemed to be telling me I was doing something wrong, which I probably was"). The software's on-screen directions did not mitigate confusion. Only one participant was able to successfully calibrate the system independently, while the others required the assistance of our designated technician.

Outcomes-Based Recommendations

Based on the testing of this complex, multi-component system situated in a complex rhetorical situation, we recommended the addition of user documentation, emphasizing the need to define key terms and to provide step-by-step instructions for use of the software and hardware. While the client responded that our testing of the hardware prototype did offer insights into that iteration of the design, it is difficult to say whether or not that feedback was ultimately helpful to the product's success because another iteration of the prototype was concurrently being built as we tested the one we had on-hand.

Usefulness of an Agile Attitude

Our service-learning course was an "agile challenge" that Barnum (as cited in Six, 2011) defined in the form of a question: "How to fit usability testing into the short timeframe of a typical sprint?" (Six, 2011, "Adapting User Research," para. 1). According to the *Manifesto for Agile Software Development* (Beck et al., 2001), the following are principal values of agile programming:

- Individuals and interactions over processes and tools
- Working software over comprehensive documentation
- Customer collaboration over contract negotiation
- Responding to change over following a plan

In discussing agile development and the user experience, Ramsay asserted "agile is not a specific process or methodology; it is a way of thinking" (Six, 2011, Introduction, para. 5). Evident in the agile way of thinking is the user-centered focus, and ultimately, it was the flexibility of the agile process that enabled us to be successful with the number of variables we were dealt. We could have easily been caught up in the complicated details of the multi-genre product we were testing, but the user-centered nature of the agile, iterative process provided us with the standpoint of seeking insight on "individuals and interactions over processes and tools" (Beck et al., 2001). Emphasis of individuals and interactions over processes and tools justified our more interactive approach to the think-aloud protocol, first encouraged by Nielsen (1993) and Tamler (1998) but later deemed unnecessary at best (Krahmer & Ummelen, 2004) and controversial or dangerous at worst (Boren & Ramey, 2000; Barnum, 2011).

The developmental status of the product created a usability-testing context focused on software functionality rather than documentation. Additionally, the educational context—living on campus for two weeks—naturally supported scrums. Sy (2006) explained that "scrums, through face-to-face communication, take the place of detailed documents to guide project planning" (p. 114).

We spent only minimal time on documentation, as the context of our sprint was changing on a daily basis and as adjustments to the software and hardware were happening on the fly. The bulk of our formal written material provided structure to and supported rigor in our testing, and even with the pressure to maintain a high degree of professionalism, that material had to be flexible to each individual user-testing instance.

Our contextual role was mixed. Because the owner and designers were constantly in close contact, the testing process was highly collaborative. However, even as we were embedded contextually in the product design process, our team also had characteristics of an external usability testing organization. We completed tasks associated with an agile testing environment: fast planning, participant recruiting, rough and tumble pilots, and quick debriefs (Six, 2011). Our "contract" for the work was verbal, and although we did provide a formal testing plan to our graduate instructor, the plan was adaptable in response to the rapidly developing nature of the product. Scrums cemented our relationships in multiple directions: among ourselves as working group members, with the product designers, with the product owner, and with our graduate professor. Especially as online students attending the in-residence program for the first time, the challenge and intensity affected "social and cultural interrelationships, instilling increased cross-disciplinary empathy and understanding [and replacing] traditional document-centered communication...with fast, lightweight communication that reflects the current project reality" (Ramey as cited in Six, 2011, "Team Dynamics," para. 4).

If we had been in a later stage of user testing for the elements of our service-learning course—hardware, software, and both interacting together—we could have perhaps met Madrigal and McClain's (2011) expectation for "testing a version or mockup that is free of glitches, bugs, or known errors" and emphasizing the importance of testing for a product in "primetime" (para. 1). We attempted to find an intermediary solution, particularly looking at lo-fi options that could draw insights and provide a less frustrating situation for the end user. However, given the number of factors that required testing and the end-goal of having a more usable product to introduce at a rapidly-approaching conference, these options were not feasible.

Conclusions

Our movement from the initial meeting with the clients until our final recommendations emphasized the rhetorically grounded nature of the testing situation and design, sometimes pushing us to stretch the boundaries of the usability scholarship we were assigned as part of the graduate course. Spinuzzi (2005) affirmed our decision to advance beyond the scope of traditional usability procedures, and claimed that much of the current (as of 2005) usability testing literature simply does not "foreground usability as a set of practices involving rhetorical translation" (p. 382). As we navigated each intersection of practice and theory, we employed techniques that were "rhetorical, context-dependent actions guided, but not wholly determined by conventional heuristics, procedures, and tools" (Spinuzzi, 2005, p. 383). Thus, despite our student anxiety over "not doing it right," our rhetorically-based choices of tasks, scenario, and script were beneficial responses to the complexity of the situation.

Johnson, Salvo, and Zoetewey (2007) described usability as bridging a gap between the scientific and cultural disciplines. As rhetoricians, we heartily agree. Our usability team's unique relation to an "embedded" client also supports Dray's (2009) call for "bridge-building" between academia and industry. Despite raising questions of bias, the relationship proved to be a valuable component of our iterative testing process. We were able to leverage this "meta-awareness" continually to inform and shape our study, as well as focus on designing rigorous methods that acknowledged this context. We had a clear sense of the clients' goals and an acute awareness of the kind of market they were targeting. Conversely, the client understood our inherent limitations as students with a significant time constraint. Dray (2009) noted that in the often time- and money-constrained business of usability, "Success can be determined not by who has the 'best' methodology or analysis, but by who best understands the concerns of the decision makers" (p. 4).

Because the product creators are members of our graduate program, their investment in the project was two-fold: (a) provide the necessary and timely background knowledge to ensure our testing could commence on schedule and yield results and (b) encourage the spirit of intellectual inquiry and academic rigor ubiquitous in our doctoral program. Likewise, although

we grew in our abilities as research practitioners over the course of two weeks, we also remained doctoral students. Therefore, within the iterative process of design, test, and redesign was another looping process at work, a thoroughly educational one. We tested a product on users that was made by users for users, and, as usability students, we were also users. In short, we designed and participated in a distinctly “meta-process,” encountering many traditional usability testing challenges, from decisions regarding participants to methodology to time constraints, while at the same time feeling added anxiety over our student-researcher role in the iterative process. Similar to the practitioner lab, our “time to think deeply about findings and interpretation [was] greatly compressed” (Dray, 2009, p. 4), yet because it was grounded firmly in an academic context, our project was intended also to yield significant intellectual growth.

Dray (2009) saw this intersection of practice and academia as a positive trend, noting that academic rigor coupled with a more academic style of critical thinking within the world of practice can provide untold gains in an increasingly competitive market. Likewise, she insisted that academia (particularly students) stands to benefit greatly from interaction with the commercial world: “[D]eeper understanding within academia of the world of practice can help academics do a better job of preparing their students to function as professionals...[and] alliances can help provide students with real-life experience” (Dray, 2009, p. 6). Indeed, our relationship with the client, the TCR program, and our classmates in the other courses supported our “studies of methodological validity under realistic conditions” (Dray, 2009, p. 6), and reinforced our collaborative efforts to achieve successful outcomes through active participation and rigorous inquiry.

As Scott (2008) described, this service-learning usability project provided “a useful set of contexts full of rhetorical tensions in which to try out usability approaches and methods” (p. 401). However, we were also accountable to a real client with a real timeline for the unveiling and marketing of a new product. In response, we approached and conducted our usability testing as both academics and practitioners. Our aim was two-fold and was embedded in a highly iterative meta-process that required both traditional and more fluid, unconventional methodologies: we sought to respect scholarly developments and strive for rigor in usability studies while making thoughtful, careful decisions where best practices did not fit our content. Hawhee’s (2004) concept of kairotic readiness, the idea of a student service-learning project that is “connected to performing” through an agile, “flexible intelligence” (p. 70, 193) is an accurate description of our experience. Consequently, our experience speaks to the value of academia encouraging students to take on real world projects as a way to not only establish closer connections with the practical world of commercial industry, but also to develop our own skills as engaged scholars.

Tips for Usability Practitioners

The following is a list of recommendations that we provide from our study:

- Consider the usefulness of rhetorical theory when unpacking complex usability situations.
- Consider the effects of complex systems, particularly where the genres and levels-of-development of multiple components complicate UX testing design.
- Prioritize an intelligent and context-specific agility over scholarly prescription when designing and conducting usability tests.
- Seek out industry-academic partnerships to promote service learning as a means of infusing industry with recent scholarly knowledge and strengthening academic experiences with real-world complexities and constraints.

Acknowledgments

We thank the faculty and staff of Texas Tech University’s graduate program in Technical Communication and Rhetoric for providing this opportunity as part of our PhD studies. Special appreciation goes to our instructor, Dr. Joyce Carter, for shepherding us through the usability literature and for allowing us the freedom to make our own path through the project. We also thank the product owners, stakeholders, and development team for their support and

participation in this experience. Finally, we are grateful to the three JUS anonymous reviewers as well as the special issue editors for their encouragement and for their valuable suggestions during the manuscript revision process.

References

- Artemeva, N. (2005). A time to speak, a time to act. *Journal of Business and Technical Communication*, 19(4), 389–421. doi:10.1177/1050651905278309
- Barnum, C. M. (2011). *Usability testing essentials*. New York: Morgan Kaufmann.
- Beck, K., Beedle M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R.C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). *Manifesto for agile software development*. Retrieved from <http://agilemanifesto.org/>
- Bitzer, L. F. (1968). The rhetorical situation. *Philosophy & Rhetoric*, 25, 1–14.
- Boren, M. T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3), 261–278.
- Caulton, D. A. (2001). Relaxing the homogeneity assumption in usability testing. *Behaviour and Information Technology*, 20(1), 1–7.
- Creswell, J. W. (1998). *Qualitative inquiry and design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
- Devitt, A. J. (2000). Integrating rhetorical and literary theories of genre. *College English*, 62, 696–717.
- Dray, S. M. (2009). Engaged scholars, thoughtful practitioners: The Interdependence of academics and practitioners in user-centered design and usability. *Journal of Usability Studies*, 5(1), 1–7.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT-press.
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35(3), 379–383.
- Hughes, M. (1999). Rigor in usability testing. *Technical Communication*, 46(4), 488–494.
- Hawhee, D. (2004). *Bodily arts: Rhetoric and athletics in ancient Greece*. Austin, TX: University of Texas Press.
- Johnson, R. R., Salvo, M. J., & Zoetewey, M. W. (2007). User-centered technology in participatory culture: Two decades “Beyond a narrow conception of usability testing.” *IEEE Transactions on Professional Communication*, 2(4), 320–332.
- Krahmer, E. & Ummelen, N. (2004). Thinking about thinking aloud: A comparison of two verbal protocols for usability testing. *IEEE Transactions on Professional Communication*, 47(2), 105–117.
- Lepore, T. (2010, May 17). Sketches and wireframes and prototypes! Oh my! Creating your own magical wizard experience. *UXmatters*. Retrieved from <http://www.uxmatters.com/mt/archives/2010/05/sketches-and-wireframes-and-prototypes-oh-my-creating-your-own-magical-wizard-experience.php>
- Madrigal, D. & McClaine B. (2011). This is not just a test: It’s primetime. *UXmatters*. Retrieved from <http://www.uxmatters.com/mt/archives/2011/05/this-is-not-just-a-testits-primetime.php>
- Makri, A., Blandford, A., & Cox, A. L. (2011). This is what I’m doing and why: Methodological reflections on a naturalistic think-aloud study of interactive information behaviour. *Information Processing and Management*, 47, 336–348.
- Nielsen, J. (1993). *Usability engineering*. Boston, MA: Academic Press, Inc.

- Nielsen, J. (2000). Why you only need to test with 5 users. Nielsen-Norman Group. Retrieved from <http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Nielsen, J. (2010). Mental models. Nielsen Norman Group. Retrieved from <http://www.nngroup.com/articles/mental-models/>
- Redish, J. (2007). Expanding usability testing to evaluate complex systems. *Journal of Usability Studies*, 2(3), 102–111.
- Redish, J. (2010). Technical communication and usability: Intertwined strands and mutual influences. *IEEE Transactions on Professional Communication*, 53(3), 191–201. doi:10.1109/TPC.2010.2052861
- Sasaki, T. (2008). Concurrent think-aloud protocol as a socially situated construct. *International Review of Applied Linguistics*, 46, 349–374.
- Scott, B. J. (2008). The practice of usability: Teaching user engagement through service-learning. *Technical Communication Quarterly*, 17(4), 381–412.
- Six, J. M. (2011). Integrating UX into agile development. *UX matters*, April 2011. Retrieved from <http://www.uxmatters.com/mt/archives/2011/04/integrating-ux-into-agile-development.php>
- Society for Technical Communication. (n.d.). Usability and user experience resources: Usability toolkit. Usability and user experience: An STC community. Retrieved from <http://www.stcsig.org/usability/resources/toolkit/toolkit.html>
- Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication*, 52, 163–174.
- Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. *CHI 2001 Extended Abstracts*, 285–286. New York, NY: ACM.
- Still, B., & Albers, M. J. (2010). Editorial: Technical communication and usability studies. *IEEE Transactions on Professional Communication*, 53(3), 189–190.
- Still, B., & Morris, J. (2010). The blank page technique: Reinvigorating paper prototyping in usability testing. *IEEE Transactions on Professional Communication*, 53(3), 144–157.
- Sutcliffe, A., Ryan, M., Doubleday, A., Springett, M. (2000). Model mismatch analysis: Towards a deeper explanation of users' usability problems. *Behaviour & Information Technology*, 19(1), 43–55.
- Sy, D. (2006). Adapting usability investigations for agile user-centered design. *Journal of Usability Studies*, 2(3), 112–132.
- Tamler, H. (1998). How (much) to intervene in a usability testing session. *Common Ground*, 8(3), 11–15.
- Yates, J., & Orlikowski, W. (2002). Genre systems: Chronos and kairos in communicative interaction. In R. M. Coe, L. Lingard, & T. Teslenko (Eds.), *The rhetoric and ideology of genre* (103–121). Cresskill, NJ: Hampton Press.
- Vatz, R. E. (1973). The myth of the rhetorical situation. *Philosophy & Rhetoric*, 6(3), 154–161.

About the Authors



Laura Cunningham

Ms. Cunningham is an administrative law judge for the Montana Department of Revenue. Her research interests include rhetorical mediation in government agencies, interdisciplinary communication practices, the ethos of policy-making, and ecofeminist rhetorics.



Russell Kirkscey

Dr. Kirkscey is a lecturer in the English Department at Texas Tech University. His research interests include health and medical rhetoric, communication ethics, rhetorical theory, and research methods.



Ashlynn Reynolds-Dyk

Ms. Reynolds-Dyk is an adjunct faculty member in supplemental studies at Colorado Mesa University and an affiliate faculty member for Colorado Christian University. She received her MA in Communication Studies from The University of Montana. Her research interests are in rhetoric, health, and disability studies.



Nancy Small

Dr. Small is an Instructional Associate Professor of English at Texas A&M University's branch campus in Doha, Qatar. Her research focuses on rhetorics of identity and persistence, particularly as constructed through narrative and metaphor.



Charity Tran

Ms. Tran is a Design Strategist for Disney Interactive Labs, specializing in design research and user experience. She received a BA in English with a minor in Web Technology and a Master of Professional Writing degree from the University of Southern California. Her research interests include rhetoric, fan studies, and new media.



Virginia Tucker

Dr. Tucker is Director of Interdisciplinary Programs at Old Dominion University and coordinator of its professional writing major. She teaches technical and scientific writing, digital writing, interdisciplinary theory, and rhetoric both on-site and online.