DANCING WITH A BEAR: ONE UNIVERSITY'S EXPERIENCE WITH EVALUATING BLENDED LEARNING

Patsy Moskal

Research Initiative for Teaching Effectiveness

University of Central Florida

ABSTRACT

The 2008 Sloan Blended Workshop evaluation track provided participants with ideas, methods, and resources with which to design and evaluate their courses, programs, and blended initiatives. The many questions that arose from participants at the Workshop provided the motivation to document our University's evaluation experiences at course, program and institutional levels. This paper presents our lessons learned including successes and challenges, with developing and conducting a longitudinal impact evaluation.

KEYWORDS

Blended Learning, Evaluation

I.   INTRODUCTION

With more than 75% of adults in the United States using the Internet [1], and 55% of those having high speed Internet connections [2], it is easy to understand why universities are turning to blended learning to better serve the needs of their students. Estimates are that nearly 55% of institutions are offering blended courses with that figure continuing to increase [3]. These courses seem to offer the best of both worlds, preserving face-to-face contact in a reduced seat time format, while allowing faculty to creatively use Web resources in instruction.

What constitutes "blended learning" varies widely in the literature. Allen and Seaman describe blended learning as a course with 30–79% of the instruction delivered online [4]. Niemiec describes UIC's blended courses as those which contain 25–74% online content [5]. Picciano [6] presents a continuum from face-to-face (with no Web components) with fully online courses at the opposite extreme. Blended courses are those in the middle containing some face-to-face and some web instruction. Unfortunately, while there are many who are incorporating blended learning into their instruction, fewer faculty and researchers evaluate their experiences in these varied instructional environments. Given the myriad definitions, technology resources and instructional styles that are being used in these courses, there is clearly a significant need for research on how, when and where these course components work.

The 2008 Sloan Blended Workshop (http://www.uic.edudepts/oce/blended/workshop/index.htm) evaluation track provided participants with ideas, methods, and resources with which to design and evaluate their courses, programs, and blended initiatives. The many questions that arose at the Workshop provided the motivation to document our evaluation experiences at course, program and institutional

levels. This paper presents our university's experiences, including successes and challenges, with developing and conducting a longitudinal impact evaluation.

II. THE UNIVERSITY OF CENTRAL FLORIDA DISTRIBUTED LEARNING IMPACT EVALUATION

The University of Central Florida (UCF) began offering fully online courses in Fall 1996 in part to provide access to students during a period in which classroom growth could not keep up with enrollment growth. The University determined that resources would have to be committed for this initiative to be successful; as a result, units were established to handle faculty development (http://cdws.ucf.edu/), marketing, registration, and scheduling (http://www.online.ucf.edu/cdl/) and research (http://rite.ucf.edu). That the university invested funding indefinitely confirms the vision of the initial administrative planners and their determination to make this work for students.

The university offered blended courses the following year (1997) after an evaluation finding that 80–90% of online students were also enrolled in face-to-face courses. The image of a "remote" student who reclusively took online courses has never occurred at UCF.

Dr. Charles Dziuban was recruited to design and conduct the evaluation of the new online course initiative. Funding was initially included for release time for Dr. Dziuban, and a doctoral student (the author) as evaluators, and a small expense budget. The Distributed Learning Impact Evaluation began with an institutional perspective centered on providing necessary assessment data to university administrators. However, as the evaluators began interacting with faculty, they found many who requested support for their classroom-based research. In a university culture that valued the Scholarship of Teaching and Learning (SoTL), providing support for faculty teaching research was added to the organic mission of this initiative. In 1999, the Research Initiative for Teaching Effectiveness (RITE) became a formal department with a modest budget that continues today. RITE personnel include the Director (Dziuban) and Associate Director (Moskal), office manager and 2–3 graduate assistants. A small travel budget also funds RITE travel and is used to assist faculty presenting at conferences.

Since its inception, RITE has been a dynamic and evolving unit, incorporating both institutional and individual classroom research perspectives. In that time, we have compiled a number of revelations and lessons learned regarding this quickly growing mode of instruction. We have found that many of these courses are also being supported by fellow researchers at other universities.

In planning the distributed learning impact evaluation, we knew that our evaluation would assist faculty on our campus who were involved with planning, design, and the development of online and blended courses. The focus from the beginning was to examine the impact these courses have on both faculty and students. As we began planning the evaluation, we developed a number of principles that have held throughout the longitudinal study:

- Evaluation must be objective.
- Evaluation should conform to the culture of the institution.
- Uncollected data cannot be analyzed.
- Data do not equal information.
- Qualitative and quantitative approaches must complement each other.
- We must show an institutional impact.

- Our results may not be generalized beyond UCF.

Evaluators have preconceived biases and ideas, some based on anecdote, as to what "results" we hope to find, but from the beginning we stressed the importance of bringing "all the news" not just the good news. In evaluating these courses, we need to know what does not work and what works well from a technology, instructional, and context standpoint to inform faculty and administrators. We also knew that our evaluation had to be designed with UCF in mind. In interacting with other university researchers, we quickly discovered that there were endless designs occurring, based on college size, support, needs, and so forth. We were not concerned with generalizing our results beyond our campus, but rather increasing quality in these courses for both faculty and students. Any generalized results were "icing on the cake" and we have made a point of interacting with researchers from other campuses to learn from them. Perhaps the most significant principle has always been to gather data whenever possible. Our motto is "you can always choose *not* to analyze data you have collected, but you can never analyze what you *do not* collect." Too often evaluation is an afterthought which makes it challenging, if not impossible, to deliver any results in a timely fashion. By working with other departments on our campus--Institutional Research, Data Mining Institute, Computer Services and Telecommunications, Faculty Center for Teaching and Learning—we have been able to collect large amounts of data and develop trends across time. We also do not rely on quantitative data exclusively, but collect qualitative information as well. While institutional datasets containing large amounts of data, such as student enrollment, grades, ratings of instruction, allow for sophisticated statistics, they do not reveal what is happening in the classroom and what works well. Through open-ended responses on student and faculty surveys, interviews and focus groups with blended participants, we often find *why* things are going well or not.

Finally, we make a point of interpreting our data and often disseminate results across our campus. Administrators rely on us to provide a snapshot analysis of the data which are presented at semester meetings with deans and in weekly lunches with directors of units involved with blended learning. However, faculty are also kept informed through new faculty orientations, summer and winter workshops, and presentations to each group who participate in the faculty development program provided to those teaching online or blended courses.

In gauging impact of online and blended courses, RITE's approach was to determine the impact on faculty and students. Figure 1 shows the design of UCF's distributed learning impact evaluation. The initial framework for the design has continued, with refinement and additions, through the years.
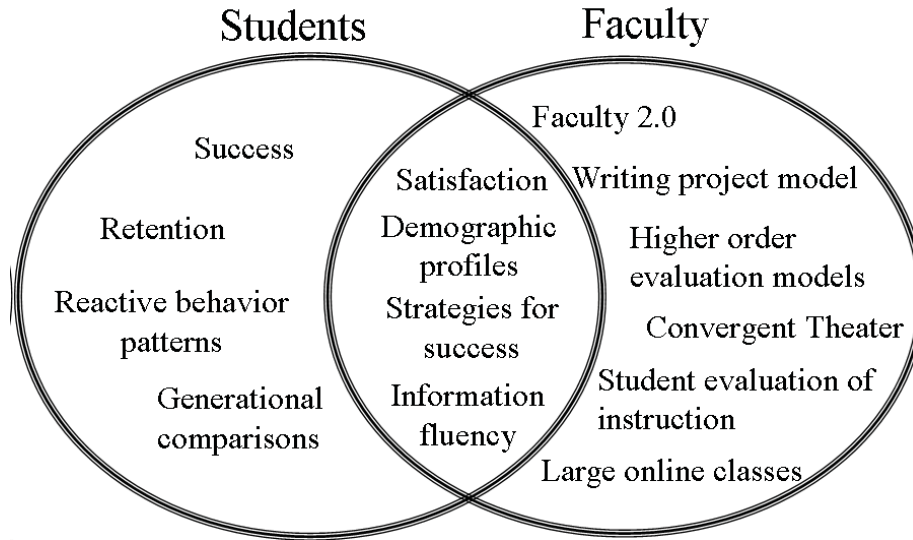
Figure 1. UCF's Distributed Learning Impact Evaluation design

A. Student Satisfaction

Through the years, RITE has administered four student surveys and has conducted student focus groups to measure student attitudes toward blended learning. Early on, surveys were mailed to students and entered into the Statistical Package for the Social Sciences (SPSS) by student assistants. Many faculty continue to use this method for their in-class surveys. As the technology became available to administer online surveys, we found a means that eliminated the need for data entry of large numbers of returned surveys. However, we have found that the online surveys (and student ratings of instruction) have lower response rates than their paper counterparts. For students (and faculty), it is too easy to ignore an e-mail request or a link on the course web page. One strategy has been to provide 'gentle reminders' to encourage students to respond. More recently we have found that targeting large classes, in which faculty are willing to provide extra credit incentives, helps us obtain more responses, albeit with a certain amount of bias depending on the course and teachers. Bias is always a concern with surveys, however. Therefore, our approach has been to gather data whenever we can. Through repetition, we have found fairly stable results. For instance, students have been consistently positive about blended courses and report that these courses make it easier for them to complete an education while juggling career, family, and other responsibilities.

An additional measure of satisfaction has been the student evaluation of instruction form, administered near the end of each semester in each course. These data are maintained and disseminated by our Computer Services and Technologies Department. Due to the size of these datasets—the current dataset we are researching contains 1.2 million records—we have worked with the UCF Data Mining Institute to develop models for what students find is "excellent" and "poor" instruction. Using the expertise of these data mining statisticians has allowed us to use more sophisticated decision tree analyses such as Classification and Regression Trees (CART). This technique allows us to identify classification "rules" for an instructor who received an Excellent, Very Good, Good, Fair, or Poor overall rating. For instance, 96% of students who indicated faculty were Excellent at facilitating learning and communicating ideas (two items on the student rating of instruction form) also gave the instructor an Excellent overall rating. Table 2 shows the percentage of Excellent ratings for each modality overall, and for those fitting this rule. We found that these models consistently hold true across departments and modalities.

Table 1. Student Ratings by Modality: Overall and for Those Rated as Excellent in
Facilitating Learning and Communicating Ideas

Course Modality	% Excellent overall	% Excellent within rule
F2F	42.0	92.2
Enhanced	44.0	92.3
Blended	40.6	92.0
Online	55.4	92.7
Interactive TV	20.9	86.7
	N=709,285	N=235,745

B. Student Success and Withdrawal

Often, the initial question that occurs regarding any change in instruction is "is it as good as…" Within the context of web courses, the standard was thought to be face-to-face courses, but on our campus there are no "pure" face-to-face courses any longer. As online and blended courses have grown, faculty have begun incorporating web components into what used to be considered face-to-face sections. Many sections are now only offered as blended, or fully online, and many others use online resources such as discussion groups, Web resources, testing, etc. From a research standpoint, locating course sections of differing modalities to compare has become impossible because Web course instruction is as varied as the instruction that occurs in what has been considered a typical face-to-face course.

Grades, therefore, constitute the only course measure that is common across all departments and colleges. However, grades and grading practices also vary widely due to instruction and assessment philosophical differences across various disciplines and faculty. To gain reliability, in lieu of comparing specific grades, we measure *success* defined as an A, B, or C grade. This tactic has provided us with a less specific, but much more reliable measure of student performance.

We also examine student success and withdrawal rates for various demographics groups: by ethnicity, gender, race, generation, department and college, for example. We collect these data from the institutional database with help from our Office of Institutional Research, carefully omitting students' personal information. Because we have examined these data since 1996, we are also able to document trends that deviate from the norm and because we now have a significant dataset we can use more sophisticated analyses such as decision trees. Figure 2 illustrates an example of using decision trees to examine these data by generation and course level.
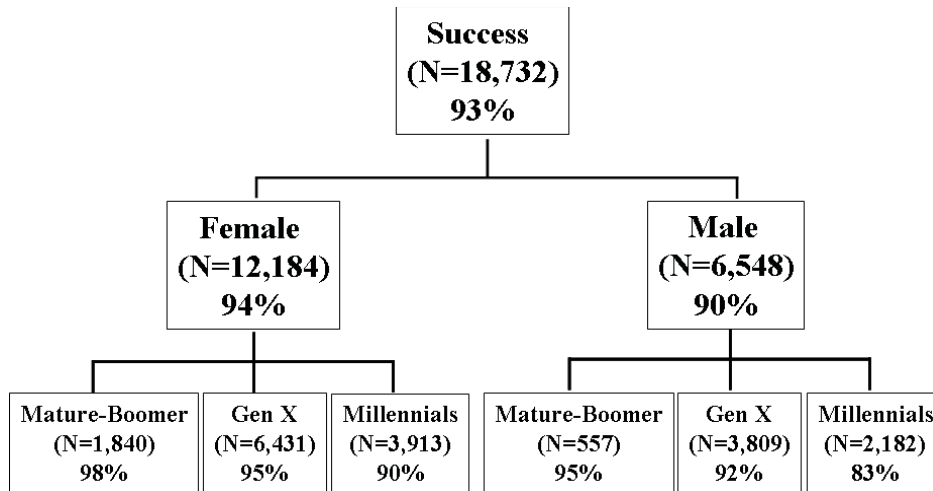
Figure 2. Student Success Rate by Gender and Generation

C. Blended Faculty

UCF's Center for Distributed Learning monitors college and department shifts in faculty teaching online and blended courses. This information is disseminated to deans and department chairs on a semester basis and used for planning purposes.

Faculty demographics and satisfaction are also routinely monitored by RITE through surveys and data collected from course and WebCT rosters of those teaching blended courses. Like students, faculty report that they like the convenience and interaction in blended courses, but also express concern with academic dishonesty and workload associated with online courses. We find that faculty are more than willing to share their successes and concerns and we provide opportunities for them to give us more than Likert feedback regarding what they like and dislike about Web teaching (Table 3). This type of information is disseminated regularly at new faculty orientations and faculty development workshops. It is also used by Course Development and Web Services in planning their training sessions for faculty new to Web teaching, and faculty are provided with strategies for dealing with challenges that may arise in this instructional environment.

Table 2. Open-Ended Comments From Faculty Describing Positives and Less Positives in Web Teaching (N=56)

Positives		Less Positives	
Flexibility of time	26%	Requires too much time	35%
Interaction with Students	26%	Lack of face-to-face contact	18%
Convenience of location	19%	Technical problems	14%
Enjoy technology	11%		

D. Faculty Research

Gathering institutional, college, and department data has served UCF well, providing support for accreditation reviews and institutional effectiveness, and for informing faculty and administrators about what works well and what does not regarding blending learning. However, from an individual faculty standpoint, the most important data involves what is occurring in his or her course. So, early on we decided to provide support for faculty to do pedagogical research in their courses. Often, faculty have

ideas for research questions, but either lack the background in statistics and research or lack the time and support to do the grunt work. We provide help for both. We help faculty work out a research design; develop or locate survey protocols, if necessary; work through the Institutional Review Board (IRB) process; help locate articles for literature review; collect and analyze data; and produce quality charts and graphs. We do not write manuscripts, but we do have an editor that can review work, and we help faculty locate appropriate journals to publish their work. All of this is free to faculty and becomes their intellectual property. The caveat is that the research must be specific to pedagogy and we caution them that if they are tenure-track to insure that this will help them in that endeavor as some must publish in discipline-specific journals for tenure and promotion.

This strategy has worked very well with faculty and resulted in many publications that might not otherwise be produced. It also allows RITE to connect faculty who are working in similar instructional areas, but might be in different disciplines; e.g., several faculty interested in using *Second Life*: one in Accounting, one in Education, and one in Chemistry.

This faculty support has led to a number of creative research projects including:
1. Using constructive engagement to improve learning by allowing students to participate in evaluation of self and peer work [7].
2. Examining the effect of paradigm of the discipline on the online learning environment [8].
3. Investigating quality assurance strategies for effective online learning [9].

Most faculty who use RITE's services need only one aspect of the research continuum and many have benefitted through presentations and publications in their disciplines. Scholarship of Teaching and Learning (SoTL) is also promoted and rewarded on our campus and faculty who do research on pedagogy are now eligible for SoTL rewards. Faculty who are interested in research to improve their instruction are the best and brightest and the academy benefits as a whole when promoting publications and presentations from the most creative faculty on campus.

III.   CONCLUSION AND LESSONS LEARNED

Since 1996, UCF's evaluation activities have evolved into today's initiative. Twelve years have provided us with many trials, errors, and lessons learned. The title of this article relates to our favorite Russian parable that we quote to others frequently: "*When you dance with a bear, you can't quit when you are tired…*" We enjoy sharing our trials and tribulations with other researchers almost as much as we enjoy learning from them. Here are some of the lessons we have learned:

Keep it simple. Our evaluation began with two people and a lot of questions and ideas. We started with a student survey and it grew from there. As online and blended courses grew on our campus, questions arose from anecdotal comments from faculty, students, instructional designers and administrators involved with online and blended learning. We listened and incorporated issues which arose into our research questions from which our evaluation design has evolved. As new questions arise, our evaluation evolves. But, starting simple was critical to our initial success. A simple, successful evaluation trumps a complicated, impossible design, every time.

Find existing data. There is a limit to available resources (namely time, money, and expertise) in every research design. We learned early on that there were data available that were collected and maintained regularly on our campus. Utilizing existing data eliminates collection time and cost. For instance, student

and faculty data are maintained by each university. On our campus, our Office of Institutional Research maintains these databases and will work with faculty or departments who need data for research. Student evaluations of instruction are also available on our campus. We work with Computer Services to collect these data as well. It would be impossible for us to gather these data on our own, but it has immensely enriched our impact evaluation.

Learn from others. In addition to working with other departments who collect data, we also look to others on our campus who have expertise that we do not. For instance, we have learned a great deal from our Data Mining Institute about working with large datasets. The field of data mining is relatively new and did not exist when we began our evaluation. After years of collecting student success and withdrawal rates, as well as student evaluations of faculty, we began looking for alternatives to analyze large datasets. Data mining has allowed us to develop useful models with these data. These techniques enhance the general statistics we continue to utilize.

There are many experienced researchers on campuses, and many who are interested in collaborating. Faculty who might not have research or statistics expertise can turn to those in departments such as Educational Research, Statistics and Psychology for help. Our office, as well as our Faculty Center for Teaching and Learning will provide assistance or connect faculty with those who can assist them.

There are also many external resources available and researchers who are eager to share their experiences. What are others doing in evaluation? What have they learned? We frequently peruse the literature for ideas on how to improve our evaluation. We freely share our survey instruments and research design and it has been our experience that other researchers are as willing to collaborate. Organizations such as Sloan-C are invaluable in providing both resources and collaborative opportunities. The Sloan-C pillars (Learning Effectiveness, Access, Cost Effectiveness, Student Satisfaction, and Faculty Satisfaction), for instance, provide an excellent model for evaluation [10].

Find the message in the data. Data by itself is not information. For your evaluation to be credible, it has to inform and have an impact. Whether it is class, program, or institution-based, we provide analyses in graphical form, clearly delineated for the stakeholders at hand. Clarity of results is essential (particularly for administrators) as your audience, whoever they may be, will have limited time and limited interest. They also may have no or minimal expertise in your area of research, or already have preexisting ideas and biases. Knowing your audience is always critical to success.

Make an impact. Design the evaluation for your course, program, initiative, and/or university. Disseminate the results to others, especially to decision makers. Evolve as the initiative evolves. A static evaluation in an evolving setting will not inform, and will quickly become obsolete. An early evaluation finding that the majority of fully online students also were enrolled in face-to-face classes contributed to the university's decision to pursue blended courses to maximize classroom space. An early finding that blended faculty were having to spend significant face-to-face time helping students with technology issues led to orientations being offered to students, a CDROM of tutorials and plug-ins, and help desk support. Through regular meetings of UCF's CIO, Directors of Course Development and Web Services, the Center for Distributed Learning, and RITE, and meetings with deans, evaluation results are used to inform and ultimately affect online and blended course practice. This iterative process has helped UCF to address issues and successes quickly and has contributed to student and faculty satisfaction (and ultimately participation) in Web courses.

Iteration is invaluable. We find that we always can improve on our effort. We pilot test surveys with colleagues, student assistants, etc., because we want to be clear and to ensure that what we are asking will answer our questions. We feel that every survey we have administered can be improved, and that any ambiguities become crystal clear as soon as we obtain all the data. Unfortunately, that is too late for the current iteration, but it certainly justifies the need to repeat data collection whenever possible. Each class of students is unique, and while we can never be certain that our results are not just an anomaly for the current sample, repeating allows us to gain confidence in the reliability of our findings.

Dancing with a bear like evaluation involves a reflective and iterative approach. Incorporating our lessons learned into each new generation of evaluation design enhances the relevance of our work for faculty, students, and the academy.

IV.   REFERENCES

1. **Madden, M.** Internet penetration and impact. Pew Internet & American Life Project. Retrieved July 11, 2008 from http://www.pewinternet.org, April 2006.
2. **Horrigan, J. B.** Home broadband adoption 2008: Adoption stalls for low-income Americans even as many broadband users opt for premium services that give them more speed. Pew Internet & American Life Project. Retrieved July 11, 2008 from http://www.pewinternet.org, 2008.
3. **Allen, I. E., J. Seaman and R. Garrett.** Blending In: The Extent and Promise of Blended Education in the United States. Needham, MA: Sloan-C, March 2007.
4. **Allen, I. E. and J. Seaman.** Online Nation: Five Years of Growth in Online Learning. Needham, MA: Sloan-C, October 2007.
5. **Niemiec, M.** Blended learning strategies, issues, barriers & considerations. Presentation at The Sloan-C Workshop on Blended Learning and Higher Education. University of Illinois at Chicago, 2008. http://www.uic.edu/depts/oee/blended/administrationtrack/Website/institutionalperspective2008.pdf.
6. **Picciano, A. G.** Blending with purpose. Presentation at The Sloan-C Workshop on Blended Learning and Higher Education. University of Illinois at Chicago, 2008.
7. **Liberman, A., K. Scharoun, T. Rotaris, M. Fottler, C. Dziuban & P. Moskal.** Teaching, learning, and the development of leadership skills through constructive engagement. *Journal of Faculty Development* 20(3): 177–186, 2005.
8. **Hornik, S., C. Saunders, Y. Li, P. Moskal & C. Dziuban.** The impact of paradigm development and course level on performance in technology-mediated learning environments. *Informing Science: The International Journal of an Emerging Transdiscipline* 11: 35–58, 2008.
9. **Lee, J. & C. Dziuban.** Using quality assurance strategies for online programs. *Educational Technology Review* 10(2): 69–78, 2002.
10. **Lorenzo, G. & J. Moore.** The Sloan Consortium report to the nation: Five pillars of quality online education. Needham, MA: Sloan-C, 2002. Retrieved July 8, 2008 from http://www.sloan-c.org/effective/pillarreport1.pdf.

V.   ABOUT THE AUTHOR

Patsy Moskal is the Associate Director for the Research Initiative for Teaching Effectiveness at the University of Central Florida (UCF) where she has been a faculty member since 1989. She received an Ed.D. from UCF specializing in Instructional Technology and Research Methods and holds BS and MS degrees in computer science. Since 1996, she has served as the liaison for faculty research of distributed learning and teaching effectiveness at UCF. Patsy specializes in statistics, graphics, program evaluation,

and applied data analysis. She has extensive experience in research methods including survey development, interviewing, and conducting focus groups and frequently serves as an evaluation consultant to school districts, and industry and government organizations. She has also received funding from several government and industrial agencies including the National Science Foundation and the Alfred P. Sloan Foundation. She has co-authored numerous articles and chapters on blended and online learning including the *Handbook of Blended Learning Environments*, *Educating the Net Generation,* and *Blended Learning: Research Perspectives*.