A Comparison of Student Ratings in Traditional and Interactive Television Courses

Morgan McCall
Livingston County (KY) Schools

Mardis Dunham *Robert Lyons*
Murray State University

Abstract

Although interactive television (ITV) allows colleges and universities to reach a wider audience, little research has been conducted exploring the effectiveness of the courses as perceived by students. This study compared student ratings of teacher effectiveness between 331 traditional courses and 125 ITV courses. The data included 456 graduate level courses over six contiguous semesters. Results clearly favored the traditional courses, followed by perceived effectiveness of the instructor at the ITV sending site and the ratings at the ITV receiving site. Implications for the use of ITV are discussed.

A Comparison of Student Ratings in Traditional and Interactive Television Courses

The rapid improvement in technology and knowledge over the past forty years has helped transform the global economy from a manufacturing base to an information base. Indeed, most jobs across the country require some post-secondary education (Carnevale, Smith, Stone, Kotamraju, Steuernagel, & Green, 2011). High schools, recognizing the necessity of having a college education in the present economy, are pushing evermore to prepare their students for college. Because of these changes, 2-year and 4-year colleges are experiencing steady growth (Bill & Melinda Gates Foundation, 2011).

Currently, increasing numbers of high school graduates are going directly into college in hopes that they will be able to find a better paying job. Other people are going to school as a way to advance in their current job or to find new, better paying positions in other fields. Due to massive lay-offs, many people now have the time to attend college. Furthermore, in his first State of the Union address, President Barack Obama announced his intention to provide a $10,000 tax credit for families and debt forgiveness for those who have been repaying student loans for over 20 years (Peterson & Staley, 2010). This has also led to an increase in enrollment.

According to Daniel (1996), the rate at which universities and colleges are being built cannot feasibly sustain the number of students enrolling in higher education. In fact, in 2009 the Obama Administration began an effort to make America more competitive by having more college graduates than any other nation (Hamilton & Babyak, 2009). They started with streamlining the financial aid process and increasing the budget for federal aid so that a greater number of Americans from low socioeconomic status households can attend post-secondary training. President Obama has also encouraged private institutions to lower their tuition and recruit more students to increase the number of college graduates (Hill, 2010).

Many public post-secondary schools, and even private, for-profit organizations like the University of Phoenix, Capella University, and Walden University, have made access to undergraduate and graduate programs available to most anyone with an internet connection. In fact, in 2008, the U.S. Department of Education National Center for Education Statistics reported that 66% of all 4200 2-year and 4-year Title IV-eligible, degree-granting institutions offered distance education courses in 2006–2007. More than 90% of all institutions with at least 3000 students offered

some type of distance education (Parsad & Lewis, 2008). The need to provide education across vast distances has prompted the implementation of audio/visual and internet technology into the realm of post-secondary education. In addition to web-based courses, Interactive Television (ITV) is being employed in an effort to provide an in-class or traditional feel to distance education. ITV is an instructional method that uses audio and visual feeds sent from one location out to several off-campus sites. This technology allows students the convenience of taking courses closer to home and has been adopted by a large number of colleges and universities (Anderson & Kent, 2002). These electronic delivery approaches enable the school and instructor to reach out to a greater number of students and, for this reason, have become especially popular at rural schools and community colleges (Mars & Ginter, 2007).

Despite the growth and popularity of distance education, little has been done to investigate the effectiveness of this method of delivery. This is disappointing given the increasing degree of accountability for public education. Given the increasing competition among post-secondary institutions, colleges are not immune to demonstrating their effectiveness. For example, accreditation bodies, such as the Southern Association of Colleges and Schools (SACS, 2004) and the National Council for Accreditation of Teacher Education (NCATE, 2008) require colleges to demonstrate their effectiveness in training their students. Typically, public schools are judged in terms of graduation rates and scores on high-stakes tests. At this level, the effectiveness of individual teachers is measured by their students' test scores. Colleges are assessed in terms of graduation rates, job placement rates, and the success of their graduates (McCaffrey et al., 2004). Instructor effectiveness at this level is measured using end-of-course teacher evaluations completed by students.

Although researchers have attempted to gain a better understanding of the method of delivery related to the student satisfaction dynamic, thus far, research on the effectiveness of distance learning approaches, such as ITV and online courses, has been limited. For example, some past research included small samples (Bland, Morrison, & Ross, 1992; Doggett, 2008) or was restricted to a specific course (Cragg, Dunning, & Ellis, 2008). This relative lack of research is surprising given the tremendous pressure commonly placed on schools to demonstrate the effectiveness of their programs. Therefore, a more comprehensive and current study with a larger sample size and a wider diversity of courses is warranted.

Student Ratings. Student ratings of instructors are used to help faculty improve their teaching by highlighting strengths and weaknesses perceived by the students (Guder & Malliaris, 2010). Despite their continued use, student ratings of instructors constantly come under negative criticism. Some believe that the questions posed on the evaluation forms cannot be answered reliably, due to the average student's knowledge of teaching (Scriven, 1995). For example, some evaluation forms may contain questions that could potentially influence the response by mentioning "extraneous and potentially prejudicial material (i.e., questions about the teacher's personality or the appeal of the subject matter)" (1995, p.2). Other questions that have led to concerns of validity and reliability include whether or not the student would recommend the course to a friend and asking the student to make subjective comparisons between instructors. Further, Scriven describes another major issue, which is that, typically, instructors receive very little guidance on how to effectively analyze, interpret, and act upon their course evaluations. However, Scriven presents arguments supporting the use of student evaluations including the fact that students

are in a unique position to judge their own increase in knowledge, comprehension, and motivation toward a subject. He also argues that students can observe and rate details pertinent to competent, organized, efficient, and enthusiastic teaching (1995).

Aleamoni (1999) pointed out that part of the lack of perceived reliability of student ratings of instructors was due to the "halo effect," whereby students rate higher those instructors who grade easier and who are personable, without consideration for the actual content, challenge of the course, and competence of the instructor. Similarly, students' locus of control can also affect their perceptions of the course and instructor (Risser, 2010). Risser reported that the greatest predictor of final exam grades and course grades was positive internal dialogue stemming from student confidence in knowing the tested material. Students with more of an internal locus tended to make higher grades (and therefore rated their instructors higher) than students with an external locus of control.

Another common concern with SRIs is the period being evaluated (Ludlow, 2005). Often, only the two semesters of the school year are considered for review. While this can still be effective in gaining information about students' perceptions of that year, it does not allow for a longitudinal look at the teacher's progress across successive years. Ludlow (2005) noted that when past performance is ignored, significant contextual information, such as patterns and consistency on evaluations over the course of successive years, is lost.

Conversely, there is substantial support for SRIs. Scriven (1995) reported that students are in a good position to judge whether tests covered the course material. Scriven also explained that students can reliably judge such instructor characteristics as their punctuality, penmanship on the board,

and other facts related to competent teaching. Similarly, Cashin (1990) claims that, more so than any other data used for faculty evaluation, SRIs tend to be statistically reliable, valid, and relatively free from bias. Data have shown that the high reliability is due to students' consistency in their instructor ratings, while the high validity is due to positive correlations between student ratings and other measures of teaching effectiveness. For example, Serdyukova, Tatum, and Serdyukova (2010) collected data from National University's Office of Institutional Research and Assessment from July 2007 to February 2008. A sample size of 32,393 evaluations were analyzed based on variables related to student self-assessment of learning, student assessment of instruction for all classes, student assessment of instruction for online classes only, student assessment of course content, and assessment of web-based technology using a 5-point Likert scale. The Cronbach's Alpha reliability index was used to calculate reliability, which yielded Alpha levels of .84 or higher (.70 or higher is considered reasonable reliability). In general, the authors concluded, "users of these ratings can be confident that they have a solid foundation for assessing teaching and learning" (p. 186).

Further, the comparative analysis by Serdyukova, Tatum, and Serdyukova (2010) produced some compelling findings. The perception that online courses receive fewer positive evaluations was confirmed, and the GPA's of students enrolled in online courses tended to be lower than those taking classes on-site. They also found that the GPA of graduate students was higher than for undergraduate classes, but that assessment ratings between the two groups were quite similar (2010). Analysis of the validity showed correlations among student satisfaction, student grades, and student ratings of their instructors. Specifically, those students with higher grades tended to be more satisfied with the courses, therefore elevating their instructors' ratings.

Effectiveness of Distance Education. Despite the popularity of distance education, relatively little research has been conducted measuring the connection between student satisfaction and their perception of the quality of distance learning courses. Frederickson, Reed, and Clifford (2005), in comparing psychology graduate students' attitudes toward online versus traditional courses at University College London, found that students were more critical of the web-based course and that students in the traditional lecture-style class were able to identify more positive aspects of their learning experience. Additionally, they found that students appreciated the peer-collaboration exercises in the online course that the others did not receive. However, it is important to note the small sample size (N=16) and short duration (12 one-hour sessions over six weeks) of this study, which limits the study's generalizability.

Research by Clow (1999) at the University of North Carolina at Pembroke provided evidence for the use of ITV courses at the graduate, but not undergraduate, level. Of the four-hundred responses obtained by Clow, 39% were from traditional courses, while 34% were from on-campus distance learning courses and 27% were from off-campus distance learning courses. Approximately 65% of the total was undergraduate courses, with 35% being graduate level. All of the courses under study were business classes. Clow used end-of-semester surveys to measure students' expectations, instructor quality, course load, students' involvement, effectiveness, and level of course demand. Clow reported that at the graduate level, students' evaluations were not affected by the ITV approach and that the value of the course, expected retention, and increased knowledge appeared to be the same regardless of format. Clow hypothesized that this was because graduate level students tend to take a more

proactive approach in the learning process and are therefore more likely to learn through any format.

Conversely, other researchers have found limitations in ITV as an effective delivery method, including lack of immediate access to instructors, dissatisfaction with the ITV instructional method, and student perception that the ITV instructor plays a less active role, making it less effective than the traditional classroom (Paulsen, Higgins, Miller, Strawser, & Boone, 1998). Their study (*N*=67) focused on students enrolled in a special education practicum course. Participants were placed either into a traditional lecture, an ITV course, or in a videotape lecture group. The study included instructor training on the technology they would be using. Data were collected from student satisfaction surveys, instructor evaluations, and achievement tests. Paulsen et al. concluded that, regardless of delivery approach, students scored equally well on quizzes and exams. They also found that students in the ITV setting were satisfied with their experience, but would have been happier in a traditional setting and that students in the ITV setting perceived their instructors as playing a less active role in the course.

Last, Anderson and Kent (2002) found, through meta-analysis, that when other variables are held constant, students rate the effectiveness of the professor's teaching lower when ITV is involved. Earlier, research by Johnson and Silvernail (1994) on the relationship between student evaluations and course satisfaction showed that satisfaction with both the course and the method of instruction were the most predictive variables for end-of-course outcomes, compared to distance and motivation. Their sample was comprised of 1,520 students enrolled in 31 ITV courses throughout the state of Maine. Like the previous studies, end-of-semester evaluations were used as the dependent variable.

In summary, the large increase in the need for alternative methods of instruction, especially distance learning

methods, has led to a need for scrutiny of end-of-semester student ratings of instructors. Past research has produced mixed results, particularly for graduate courses, with some studies finding that the distance learning approach is effective, and others suggesting that a lack of standardized procedures inhibits its usefulness, which leads to instruction that is unequal across settings. A more in-depth investigation, using a larger sample size ranging a wide variety of graduate courses, was warranted.

Method

Procedures

Student evaluations in the form of the student completed Instructional Assessment System (IAS) teacher ratings were obtained from files in the college of education dean's office. The college of education under study was a branch of a university of 11,000 students in the southeastern United States. The first four items and a summary item (for items 1 through 4) from course evaluations from all instructors from Fall 2007 to Spring 2010 were included in the study. The rationale for using only these items is that they were the only items that were identical across the various forms offered by the IAS. The raw data (in the form of median scores) were entered into an Excel spreadsheet and ultimately uploaded to SPSS for subsequent analyses. All data were anonymous—no identifying data were collected. For this study, student ratings of the instructor's effectiveness were the dependent variable, while the course delivery method was the independent variable. All courses in the study were offered at night and taken by graduate students pursuing a master's degree or graduate coursework for continuing teacher or counselor certification.

The three course delivery methods were traditional, ITV sending site, and ITV receiving site. The traditional

delivery method describes those face-to-face courses where the instructor was physically present in the classroom with students (typically 12 to 18). Conversely, the ITV courses have a sending site or "origination site" and a receiving site—the sending site delivery method were those classes where the instructor is physically present in a classroom, with 4 or 5 students, but the instruction is delivered via ITV to four receiving sites. The sending and receiving sites had very similar number of students (3 or 4 per sending site and each receiving site). For both the ITV sending site and receiving site, students are capable of interacting (visually and verbally) with the instructor and other students, and the instructor can monitor students via a screen.

Participants

There were 54 instructors (33 full time instructors and 21 adjunct instructors) teaching 456 graduate courses across six academic semesters. Proportionally, full time and adjunct instructors taught the same number of ITV and traditional courses. Similarly, the teaching load was equally dispersed across ranks—professors taught as many courses each semester as assistant and associate professors. Summer courses were not included as teacher ratings are not obtained for summer school courses. Similarly, web-based courses were not included since reliable course evaluations are seldom obtained. There was a wide range of graduate courses included in this study, including teacher leader, general and early childhood education, school and mental health counseling, school administration, special education, and school psychology.

Results

It was hypothesized that student ratings for the traditional courses would be significantly higher than those for both the ITV sending site and the ITV receiving site, and that the

sending site ratings would be higher than those at the receiving site. Before testing this hypothesis, group means for each of the survey items and the combined score were calculated. Table 1 describes the mean scores for each of the survey items and for the course delivery.

Table 1
Mean Student Ratings Data for the Course Delivery Methods

Dependent Variables	Traditional		Sending Site		Receiving Site	
	M	*SD*	*M*	*SD*	*M*	*SD*
Question 1	4.05	.60	3.90	.84	3.55	.86
Question 2	4.03	.62	3.90	.81	3.62	.81
Question 3	4.20	.59	4.03	.81	3.73	.90
Question 4	4.11	.65	4.03	.81	3.62	.95
Combined 1 – 4	4.10	.60	3.97	.81	3.64	.85

Total *N* = 456
***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106**

These mean scores were generally consistent with the average instructor rating for this college of education, which was 4.00 out of 5.00 (*SD*=0.8) (unpublished data).

Next, a MANOVA was calculated to determine if student ratings on the four items and the combined item differed by instructional delivery method. A MANOVA is appropriate to test the hypothesis that two or more groups differ on two or more normally distributed dependent variables. In this case, the groups, or independent variables, were course delivery method. The dependent variables were the first four items of the IAS instrument, as well as a combination of items 1 through 4. An important assumption when using MANOVA is that the number of observations in each cell is equal. If this assumption is violated, the alpha

levels (the chance of finding a statistically significant finding) can be distorted, which in turn can distort interpretation of the findings. As expected given the unequal sample sizes, the Box's *M* test was statistically significant (Box's $M = 299.25$, $F(30, 8016) = 9.35$, $p = .000$), meaning that the assumption of equality of within-group covariance was violated. To compensate for this violation, Pillai's Trace statistic was employed to test the significance of the MANOVA. The Pillai's Trace criterion is considered powerful and robust and is used to compensate for potential distortion seen when Box's *M* is statistically significant. The results of the one-way MANOVA was significant (Pillai's Trace = .186, $F = 9.205$, $p = .000$) and post hoc one-way analyses of variance were computed in order to discover the differences among course delivery method and student ratings. Statistically significant differences were found across all five dependent variables. Table 2 summarizes these statistical results.

Table 2
Tests for Between-Subject Effects

Dependent Variable	Sum of Squares	*df*	Mean Square	*F*	*p*
Question 1	39.818	2	19.909	42.954	.000
Question 2	28.725	2	14.363	31.309	.000
Question 3	42.913	2	21.457	45.925	.000
Question 4	45.891	2	22.946	41.758	.000
Combined 1 – 4	38.963	2	19.482	42.749	.000

Total *N* = 456 courses

Finally, post-hoc comparisons using the Tamhane statistic were conducted to determine the presence of any significant pair-wise differences. Tamhane is appropriate when the sample sizes and variances are unequal across groups. For all comparisons, the more conservative .01 level of significance was chosen. First, the Traditional delivery

method was statistically equal to the sending site on all four questions and on the combined question. Second, the sending site was statistically equal to the receiving site across the five dependent variables. Lastly, the Traditional method was statistically higher than the receiving site on all five dependent variables. Effect size calculations using Cohen's *d* were large, ranging from .817 to .959. Essentially, any Cohen's *d* above .8 is considered large. A summary of these findings are provided in tables 3 through 7.

Table 3
Post hoc Comparisons for Question 1: The Course as a Whole

Delivery Method	Mean 1	Mean 2	Mean Diff.	Stand. Error	*p*	*d*
Traditional vs Send.	4.05	3.90	.382	.200	.186	
Traditional vs Rec.	4.05	3.55	.697	.090	.000*	.94
Sending vs Rec.	3.90	3.55	.315	.210	.379	

***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106; * = significant at .01 or less**

Table 4

Post hoc Comparisons for Question 2: The Course Content

Delivery Method	Mean 1	Mean 2	Mean Diff.	Stand. Error	*p*	*d*
Traditional vs Send.	4.03	3.90	.349	.190	.226	
Traditional vs Rec.	4.05	3.62	.590	.090	.000*	.817
Sending vs Rec.	3.90	3.62	.242	.203	.569	

***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106; * = significant at .01 or less**

Table 5
Post hoc Comparisons for Question 3: The Instructor's Contribution to the Course

Delivery Method	Mean 1	Mean 2	Mean Diff.	Stand. Error	*p*	*d*
Traditional vs Send.	4.20	4.03	.402	.189	.133	
Traditional vs Rec.	4.20	3.73	.723	.093	.000*	.959
Sending vs Rec.	4.03	3.73	.321	.205	.342	

***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106; * = significant at .01 or less**

Table 6
Post hoc Comparisons for Question 4: The Instructor's Effectiveness in Teaching Subject Matter

Delivery Method	Mean 1	Mean 2	Mean Diff.	Stand. Error	*p*	*d*
Traditional vs Send.	4.11	4.03	.346	.210	.307	
Traditional vs Rec.	4.11	3.62	.752	.099	.000*	.926
Sending vs Rec.	4.03	3.62	.407	.226	.231	

***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106; * = significant at .01 or less**

Table 7
Post hoc Comparisons for Combined Questions 1 – 4

Delivery Method	Mean 1	Mean 2	Mean Diff.	Stand. Error	*p*	*d*
Traditional vs Send.	4.10	3.97	.371	.190	.182	
Traditional vs Rec.	4.10	3.64	.690	.090	.000*	.936
Sending vs Rec.	3.97	3.64	.318	.204	.345	

***N* Traditional = 331; *N* Sending = 19; *N* Receiving = 106; * = significant at .01 or less**

As noted in the tables, an effect size was only calculated when differences between the two groups were statistically significant. In summary, these statistical comparisons revealed that the traditional course delivery method was consistently rated as significantly higher by students than the ITV receiving site. However, the ITV sending site was statistically equal to both the traditional method and the receiving site.

Discussion

It is clear that students rate the instructor significantly lower in the ITV courses. These findings are consistent with those described by others (Anderson & Kent, 2002; Frederickson, Reed, & Clifford, 2005; Paulsen, Higgins, Miller, Strawser, & Boone, 1998) who found that distance learning courses were rated significantly lower than traditional courses, but they are inconsistent with the findings of Clow (1999) who found that ratings between ITV and traditional courses at the graduate level were equal. The current study addressed the limitations of previous research studies by including a large sample size, a wider range of graduate courses (all disciplines within this college of education), and a wider range of semesters (10 semesters). Thus, this study attempted to address many of the limitations found in other studies. Although there were statistically significant differences in favor of the traditional courses, the reason for these results are unclear. Significant differences in ratings between Traditional and Receiving sites and Sending and Receiving sites could be attributed to a number of factors. One possibility is that the ITV instructors had not been properly trained in implementing this delivery system. Another is that the technology being used to transmit the audio/video was not working as intended, creating technical issues that may be seen as problematic by students and therefore attributed to the instructor. In addition, inaccurate student expectations of the ITV delivery system could attribute to differences.

Obviously, ITV is not going away as a course delivery method, nor should it, as ITV provides the opportunity for colleges to reach a much broader audience and to serve a larger, more diverse student population. However, it is clear that this delivery method needs to be modified in order to address the problems found in this research. First, untenured

and unpromoted faculty (and their superiors) need to understand that ITV course ratings will very likely be lower than those from traditional courses. Indeed, it would be unfair to equate ITV and traditional course evaluations. In light of these findings, a sliding scale should be considered, or perhaps untenured faculty should be limited to only one ITV course per academic year to prevent distortion of their overall teaching effectiveness. Second, those faculty using ITV may benefit from additional instruction on how to use the ITV technology and how to engage students most effectively. Perhaps, for example, it would be beneficial to rotate the sending site weekly. Third, students should be informed of the limitations inherent in this course delivery method. For example, ITV is not equivalent to a traditional course, student engagement cannot be monitored as closely, and instructors cannot be as responsive to questions. Any technical problems need to be addressed, as these may be attributed to the instructor, thus lowering their course evaluations. Lastly, it may be helpful to determine which faculty consistently have the highest ITV ratings and ask them to describe their teaching methods to those ITV instructors who generate lower student ratings.

Limitations. As with any study, there are limitations that hinder the applicability and utility of the findings. First, and most notably, is the lack of independence for the instructors and students. The instructors, for example, taught more than one course each semester and many students probably took more than one course per semester. The data needed to verify the extent of this dependence was not available. However, this lack of independence would likely increase the likelihood of finding a statistically significant difference among/between the groups. Future studies of this nature should consider this lack of statistical independence and address the issue through hierarchical level modeling. Second, evaluations from one

university's college of education were used for this study. Consequently, these results may not readily generalize to other universities or other programs. Third, class size was not controlled—there was no way of determining if class size influenced the results. However, it should be noted that class sizes were similar. Lastly, courses were not disaggregated; there was no way to determine if certain courses, programs, or faculty tended to generate the lowest (or highest) ratings.

Future Research. Several important aspects of this study need to be investigated further. First, future research into the effectiveness of student ratings of instruction should include programs outside of education. Second, combining evaluation data from multiple colleges/universities and disaggregating courses by program and class size should provide insight regarding the effectiveness of ITV and promote generalization. Third, researchers should ensure that all receiving sites have equitable equipment. Fourth, future research should control for the number of receiving sites. For example, some of the ITV courses in the current study had only two or three receiving sites, while other courses had five or six. Lastly, comparisons using differences between ITV and traditional delivery methods should be conducted by holding the instructor, size of the class, and the course constant.

References

Aleamoni, L. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*(2), 153-166.

Anderson, L., & Kent, C. (2002). Interactive televised courses: Student perceptions of teaching effectiveness, with recommendations. *College Teaching, 50*(2), 67-74.

Bill & Melinda Gates Foundation. (2011). *Postsecondary education.* Retrieved December 5, 2011, from http://www.gatesfoundation.org/postsecondaryeducation/Pages/default.aspx

Bland, K., Morrison, G., & Ross, S. (1992). Student attitudes toward learning link: A distance education project. *Annual Mid-South Educational Research Association* (November).

Carnevale, A., Smith, N., Stone, J., Kotamraju, P., Steuernagel, B., & Green, K. (2011). *Career clusters: Forecasting demand for high school through college jobs, 2008-2018*. Retrieved December 5, 2011 from http://www9.georgetown.edu/grad/gppi/hpi/cew/pdfs/clusters-complete-update1.pdf

Cashin, W. E. (1990). Student ratings of teaching: A summary of the research. *Management Newsletter, 4*(1), 2-7.

Clow, K. (1999). Interactive distance learning: Impact on student course evaluations. *Journal of Marketing Education, 21*(2), 97-105.

Cragg, C., Dunning, J., & Ellis, J. (2008). Teacher and student behaviors in face-to-face and on-line courses: Dealing with complex concepts. *Journal of Distance Education 22*(3), 115-128.

Daniel, J. (1996). Mega-universities and knowledge media: Technology strategies for higher education. London: Kogan Page.

Doggett, A. (2008). The videoconferencing classroom: What do students think? *Journal of Industrial Education 44*(4), 29-41.

Frederickson, N., Reed, P., & Clifford, V. (2005). Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education*, *50*(4), 645-64.

Guder, F., & Malliaris, M. (2010). Online and paper course evaluations. *American Journal of Business Education, 3*(2), 131-137.

Hamilton, J., & Babyak, S. (2009). Obama administration announces streamlined college aid application. *U.S. Department of Education.* Retrieved October 13, 2011, from http://www2.ed.gov/news/pressreleases/2009 /06/ 06242009.html

Hill, E. (2010). Obama striving for increased college enrollment by 2020. *The Pendulum.* Retrieved October 13, 2011, from http://www.elon.edu/pendulum/ Story.aspx?id=3251

Johnson, J., & Silvernail, D. (1994). Impact of interactive television and distance education on student evaluation of courses: A causal model. *Community College Journal of Research and Practice, 18,* 431-440.

Ludlow, L. (2005). A longitudinal approach to understanding course evaluations. *Practical Assessment, Research and Evaluation 10*(1). Available online: http://pareoneline. net/getvn.asp/v=10&n=1

Mars, M., & Ginter, M. (2007). Connecting organizational environments with the instructional technology practices of community college faculty. *Community College Review, 34*(4), 324-343.

McCaffrey, D., Lockwood, J., Koretz, D., Louis, K., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*(1), 67-101.

National Council for Accreditation of Teacher Education. *Unit standards in effect 2008*. Retrieved December 3, 2011, from http://www.ncate.org/Standards/ NCATEUnitStandards/UnitStandardsinEffect2008/t abid/476/Default.aspx

Parsad, B., & Lewis, L. (2008). *Distance Education of degree-granting postsecondary institutions: 2006-2007* (NCES 2009-044). National Center for Education Statistics, Institute of Education Sciences, U.S., Department of Education. Washington, DC.

Paulsen, K., Higgins, K., Miller, S., Strawser, S., & Boone, R. (1998). Delivering instruction via interactive television and videotape: Student achievement and satisfaction. *Journal of Special Education, 13(4),* 58-77.

Peterson, M., & Staley, O. (2010). Obama pushes college tax credit and debt forgiveness (update 1). *Bloomberg.* Retrieved October 13, 2011, from http://www.bloomberg.com/apps/news?pid=newsarchive&sid=acv.N66V9Bl8

Risser, H. (2010). Internal and external comments on course evaluations and their relationship to course grades. *The Montana Mathematics Enthusiast, 7*(2 & 3), 401-412.

Scriven, M. (1995). Student ratings offer useful input to teacher evaluations. *Practical Assessment, Research, & Evaluation, 4*(7). Retrieved August 11, 2010, from http://PAREonline.net/getvn.asp?v=4&n=7

Serdyukova, N., Tatum, B., & Serdyukova, P. (2010). Student evaluations of courses and teachers. *Journal of Research in Innovative Teaching, 3*(1), 173-182.

Southern Association of Colleges and Schools Commission on Colleges. *Principles of Accreditation: Foundations for Quality Enhancement.* Decatur, GA: 2004.