

The "No Significant Difference" Phenomenon: A Literature Review

Thomas R. Ramage
Associate Vice President
Academic Services
Parkland College
2400 W. Bradley Ave.
Champaign, Il 61821
Email: ramage@parkland.cc.il.us

Introduction

A quick look at the "No Significant Difference Phenomenon" website might lead the casual observer to the conclusion that an overwhelming amount of data exists to support the notion that technologically-mediated instruction and/or distance education, in nearly every form imaginable, has proven to be an effective and sometimes preferred method of educating students outside the confines of what is commonly referred to as the "traditional classroom"(Thomas Russell, 2001). From 1928 to the present, Russell has cataloged at least 355 studies, technical reports, and dissertations that have reviewed student learning outcomes in the form of satisfaction surveys, grade comparisons, standardized test scores, common embedded questions, frequency of interaction between students and faculty, and a dizzying array of other "measures" ostensibly aimed at determining if any measurable or statistically significant differences exist.

At face value, it seems that comparison or outcomes studies would be one of the most effective methods for determining the effectiveness of various educational technologies. Since the 1994 publication of Richard Clark's famous statement cautioning educational researchers to "give up your enthusiasm for the belief that media attributes cause learning", he has convinced many researchers in the field that most if not all of "No Significant Difference" studies were in some way flawed. These studies inadvertently had attributed outcomes to differences in media rather than method (Richard Clark, 1994, p. 28). Simply stated, Clark presents the idea that measurable learner outcomes, when replicable using different media, indicate that the selection of the media has little to do with learner outcomes, rather the method that the media share in delivering content is the true catalyst that leads to understanding. Further, Clark stated that, "there are no benefits to be gained from employing different media in instruction"(Richard Clark, 1983, p. 450). Based on Clark's thinking, it would seem that the 355 reports contained in Russell's "No Significant Difference Phenomenon" website have focused primarily on differences in the media rather than the methods employed *via* the medium.

This literature review will focus on the body of research available specifically related to online or Internet-based delivery systems, when compared with Richard Clark's theories, versus a sampling of research supporting the "No Significant Difference" concept.

Methodology

For comparison purposes, several recent studies related to the efficiency, effectiveness, and appropriateness of online instruction were reviewed. The term "online instruction" is a generic term that, by definition, implies that the student is physically separated from the faculty and "connected" through

the use of a computer and a network or Internet link. Related terms include Computer Mediated Conferencing (CMC), Computer Mediated Instruction (CMI), Computer Assisted Instruction (CAI), and if the element of time is taken into account, Asynchronous Learning Networks (ALN). All of these terms, for purposes of this review, are synonymous, given that the differences in each definition fall within the context of the studies used as source material in this review. The instructional (media) elements employed within the context of this definition of online instruction could consist of text, video, audio, graphics, or any combination of elements and are not differentiated. Traditional classroom instruction is defined as time and place bound, face-to-face instruction, typically conducted in an educational setting and consisting primarily of a lecture/note-taking model.

Included in this review are studies that evaluated programs or courses in which content and communication were delivered primarily via online instruction, were no older than 1994, were limited exclusively to higher education, and focused on comparing student outcomes with traditional classroom-based counterparts.

Additionally, several writings by Richard Clark, Michael Moore, Stephen Ehrmann, David Diaz, Farhad Saba, Robert Kozma, and Tom Cobb were used as a *baseline* to determine this author's perception of the consequences, reliability, and validity of the *effectiveness studies*. Additional supportive research is included to provide a framework around which *baseline* and *effectiveness* studies are placed into context.

Most sources used in this review are primary sources retrieved from peer-mediated, academic journals with the exception of the "No Significant Difference Phenomenon" website and several (noted) institutional studies retrieved from campus websites. Many of the sources employed were found in the bibliographic materials cited in the *baseline* writings. The balance of the sources referenced were retrieved from books and journals discovered through electronic databases and the author's personal library.

Analysis

Effectiveness Studies/Grades

A multitude of studies have been conducted that review effectiveness of online instruction when compared with the traditional classroom, focusing on measurable student outcomes. Several have focused on end-of-course grades (Smeaton & Keogh, 1999, Wade, 1999, Lin & Davidson, 1994, Sener & Stover, 2000, Navarro & Shoemaker, 1999, Schulman & Sims, 1999), citing either no differences to as much as 5% increases over their traditional counterparts. One notable study attempted to remove perceived instructor-bias by blind-scoring tests in a graduate-level online/traditional course, finding, "...average score for the online class was 5 points (5%) higher than for the on campus class." (Fallah & Ubell, 2000).

Interestingly, the author found no studies that exposed *lower* grades or test scores of online students compared to traditional students.

Faculty Perceptions

In a survey of 250 teachers in 1999, "... 85% of faculty felt that student learning outcomes in online education were *comparable* or better to those found in face-to-face classrooms. While this was consistent with much of the data that exists in support of the 'No Significant Difference Phenomena', it was encouraging to see that so many faculty (with experience in teaching the same course in face-to-face and

online environments) were in support of online courses being as *effective* as classroom courses." (Dobrin, 1999). Another study conducted by North Carolina State University examined outcomes of 25 new online courses, entitled Project 25. Some findings from the study include faculty reporting "high-stress" and "large time investments while learning technologies employed in online course delivery". In general, most faculty felt "online course delivery was as effective as traditional means, but expressed concerns related to the promotion and encouragement of interaction with students" (Hoey, Pettitt, Brawner, & Mull, 1997).

Student Preceptions

Johnstone, Zuniga, and Markwood (1994, p. 135) found that concerning student's feelings regarding five different technologies used to deliver coursework, more than half responded with strongly positive attitudes toward the use of technology, with computer programs and e-mail ranking the highest with 90% and 74% high-positive rankings, respectively.

Kearsley, Lynch, & Wizer (1995, p. 37), upon review of six studies which examined effectiveness and impact of computer conferencing in graduate education, drew the inference that, "the overwhelming conclusion from these studies is that computer conferencing works very well for graduate level education". In order to arrive at this conclusion, they cite higher student satisfaction, equivalent or better grade point average, higher levels of critical thinking and problem-solving skills, and more discussion between students and instructors.

Not surprisingly, much research has shown that students who self-select into online courses tend to be independent learners and prefer conceptual learning styles. Students that prefer traditional classrooms tend to be more dependent, exhibiting a preference for social and applied learning styles (Diaz & Cartnal, 1999, p. 132). Other research has shown that when asked, students prefer online courses for reasons of "cost convenience and flexibility". Conversely, student's reasons for choosing a traditional classroom included "the perception of increased opportunity for interaction, decreased opportunity to procrastinate, immediate feedback, and more meaningful learning activities". This study also showed no significant difference in examination scores for three exams and final course grade (Leasure, Davis, & Thievon, 2000, p. 149).

Baseline Writings

The *effectiveness* studies cited above provide a compelling picture surrounding the efficacy of online education. Proponents of Richard Clark's writings (Clark, R. 1983), (Clark, R. 1994) could and have applied the medium vs. method tests and for the most part, discounted each and every instance of the use of a particular media (CMI,CMC,CAI,& ALN) as non-causal factors in determining effectiveness. Clark would argue that good instructional methods, regardless of the software and hardware choices made by instructional designers, had no impact whatsoever on learning outcomes. The same outcomes could be replicated through nearly any appropriate media. Clark refers to the word "efficiency" several times, attributing it only to the selection of media for purposes of expediency. Cobb (1997, p. 21) introduces the idea of "cognitive efficiency" as a counter to Clark's theories, contending that media selection does have a relationship to learning outcomes by using an effective example. If a student is attempting to learn to identify differences in birdcalls, he has several different media choices. An instructional designer may choose to describe differences verbally, express the calls through musical notation or phonetic sounds, or simply record an actual birdcall. All of these examples *could* adequately teach a student to differentiate birdcalls objectively in an assessment, provided the student understood musical notation, phonetics, and

could hear. It would seem that the most efficient media would be the latter, most efficient media and method to teach this example. Kozma (1994) supports the idea that the media is not irrelevant and that the choices made regarding the technology selection are indeed as important as instructional methods. Efficiency plays a critical role. Although Cobb's "cognitive efficiency" proposition does not totally discount Clark's position, it does bring the role of the instructional designer into a more prominent position with respect to the need for appropriate selection of media and the impact it can have on learner outcomes. Kozma advises further research on determining which technologies are best suited to complement the best methods. The method and the media are important.

In a 1991 review of research related to the correlation of grades to work achievement post graduation, Pascarella and Terenzini (1991) reviewed over 2,600 studies on the effects of college on a variety of components, notably finding that a high quality undergraduate experience is a function of both academic performance and satisfying social relations. Additionally, they reviewed a significant number of studies that the correlation between student transcripts and work achievement is nearly irrelevant. This is significant in that it calls into question the validity and reliability of not only effectiveness studies related to technology and education, but traditional educational methods as well. Ehrmann (1995 p. 22) picks up on this thread in order to make the point that it must first be determined that what is being taught is appropriate, before any consideration is given to technology (media). Ehrmann further questions the reliability of grading systems as predictors of success when compared with Pascarella & Terenzini's report that there is no relationship between the results of the objective measure (grades) of educational experiences and the impact on the products (work experiences) that they are designed to impact most directly.

Moore and Thompson (1997 p. 59) concluded after a literature review that effective distance education is "...measured by the achievement of learning, by the attitudes of students and teachers, and by return on investment". They also pointed to problems in the design of many research studies, including some of those cited earlier as *effectiveness studies*, "specifically in regard to control of the populations being compared,...the treatments being given, and the statistical techniques being applied"(Moore & Thompson, 1997, p. 59).

Phipps & Merisotis (1999, p. 3-4) point to four key shortcomings in the research: lack of control of extraneous variables, lack of randomness in selection of subjects, validity and reliability of instruments, and the under-weighting of the importance of student and faculty perceptions.

It is the author's opinion that Phipps & Merisotis, as well as Russell, make individual studies indefensible, by lumping all research into effectiveness of technology in education into one over-arching "no significant difference" bucket and over-generalizing when evaluating the research. There are substantial differences in the design, use, evaluation, and study of media. For example, the study of interactive compressed-digital television has variables inherent to the media that may not be present in a non-interactive media. At minimum, the concept of "noise", defined broadly as anything that interferes with the audience's ability to receive a message, has a higher likelihood of manifestation in an interactive media. Simply stated, there are more things that can go wrong. Yet, reviews of the research on both sides of the effectiveness argument tend to lump disparate studies into the same mold in order to serve their own ends.

Phipps & Merisotis stated that, "...much of the research is to ascertain how technology affects student learning and student satisfaction, many of the results seem to indicate that technology is not nearly as

important as other factors, such as learning tasks, learner characteristics, student motivation, and the instructor" (Phipps & Merisotis, 1999, p. 31). This Clarkian argument is countered by Saba (1999, p. 2), explaining that, "...the proper question is not whether distance education is comparable to a hypothetical "traditional," or face-to-face, instruction, but if there is enough interaction between the learner and the instructor for the learner to find meaning and develop new knowledge". Gary Brown and Mary Wack put forth an interesting response to the critique of *effectiveness* by asking if "compelling evidence is attainable, and, second, that, even amid "dizzying" technological change and shifting student populations, such comparisons with conventional education are relevant"(Brown, G. & Wack, M., 1999, p. 2).

Conclusion

There are as many definitions of distance education as there are terms to describe distance education. The fact that distance education has been around in a myriad of forms since the first spoken word was uttered, makes explaining what it is and what it is not a difficult, convoluted, and by no means, quick endeavor. Researchers have looked at distance education from many perspectives, excluding and including various technologies from definitions, in order to make a definitive pronouncement in response to what would seem to be a very simple question: Does technology impact learning? This review found no study, no evidence of any kind that categorically proves that technology *does not* impact learning in some way, positively or negatively. There is great debate, with substantial arguments and supporting data on both sides, and many suggestions from those doing the research on how it might be improved. Why then, can't a definitive study be constructed that contains the salient variables, identifies and accounts for gaps in past research, and come up with a measurable difference one way or the other?

Running the risk of over-simplification, the reasons why this goal might never be reached are obvious.

The same research gaps outlined by Phipps & Merisotis with respect to effectiveness exist in the traditional classroom as well. Consider learning styles, multiple intelligences, socio-economic factors influencing learning and cognition, student preferences, faculty teaching style, active-learning techniques, constructivism vs. behaviorism vs. instructivist methods, air temperature, room size, time of day, and daily caloric intake. Every one of these variables has been studied to some extent and when assessed individually, some generalizations can be made, but always with exception and never with absolute certainty. Why should the study of distance education or the effects of technology on learning be held to a higher standard? "There is as much difference between two teachers doing, purportedly, the same thing in conventional classes as there is between two teachers doing different things" (Worthen, et al., 1997 as cited in Brown & Wack, 1999). It is difficult, if not impossible, to apply scientific methods to social science hypothesis. Human cognition has, to date, provided no quantifiable absolutes or baseline from which research can benchmark. At the moment, the best that can be done is to try - by constructing studies and research that address both sides of the efficiency question. Studies need to review the impact of media and method, account for efficiency of design and cognitive efficiency, and to ensure that the right questions are asked and the right messages are taught.

References

Brown, G. & Wack, M. (1999). The difference frenzy and matching buckshot with buckshot. Technology Source, May/June.

Clark, R. (1983). Reconsidering research on learning from media. Review of Educational Research, 53(4), 445-459.

- Clark, R. (1994). Media will never influence learning. Educational Technology Research and Development, 42(2), 21-29.
- Cobb, T. (1997). Cognitive efficiency: toward a revised theory of media. Educational Technology, Research and Development, 45(4), 21-35.
- Diaz, D., Cartnal, R. (1999). Students' learning styles in two classes: online distance learning and equivalent on-campus. College Teaching, 47(4), 130-135.
- Dobrin, J. (1999). Who's teaching online. ITPE News, 2(12), 6-7.
- Ehrmann, S.C. (1995). Asking the right questions. Change, 27 Mr/Ap, 20-27.
- Fallah, M.H.& Ubell, R. (2000). Blind scores in a graduate test: conventional compared with web-based outcomes. ALN Magazine, 4(2).
- Hoey, J., Pettitt, J., Brawner, C., & Mull, S. (1997). Project 25: first semester assessment: A report on the implementation of courses offered on the Internet as part of Project 25 in the fall semester. Retrieved July 30, 2001 from the World Wide Web: <http://assessment.gmu.edu/reports/Eng302/>
- Johnstone, S., Zuniga, R., Markwood, R. (1994). Student and faculty survey responses and analysis. In R. Markwood & S. Johnstone (Eds.), New pathways to a Degree: Technology Opens the College (1st ed., pp. 131-147). Boulder, CO: Western Interstate Commission on Higher Education.
- Kearsley, G., Lynch, W., & Wizer, D. (1995). The effectiveness and impact of computer conferencing in graduate education. Educational Technology, 35(6), 37-42.
- Kozma, R. (1994). Will media influence learning? Reframing the debate. Educational Technology Research and Development, 42(2), 7-19.
- Leasure, A.R., Davis, L., Thievon, S.L. (2000). Comparison of student outcomes and preferences in a traditional vs. world wide web-based baccalaureate nursing research course. Journal of Nursing Education, 39(4), 149-154.
- Lin, C., Davidson, G. (1994). Effects of linking structure and cognitive style on students' performance and attitude in a computer-based hypertext environment. In *Proceedings: Selected Research and Development Presentations at the National Convention of the AECT, Nashville, 1994*. AECT 451-461.
- Moore, M.G., Thompson, M.M. (1997). The effects of distance learning Rev. ed. (ACSDE Research monograph #15). University Park, PA: American Center for the Study of Distance Education, Pennsylvania State University.
- Navarro, P., Shoemaker, J. (1999). The power of cyberlearning: An empirical test. Journal of Computing in Higher Education, 11(1), 33.
- Pascarella, E.T., Terenzini, P.T. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco: Jossey-Bass.
- Phipps, R., Merisotis, J. (1999). What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education (). Washington DC: The Institute for Higher Education Policy.

- Russell, T. (2001). The "No Significant Difference Phenomenon". Retrieved July 23, 2001 from the World Wide Web: <http://teleeducation.nb.ca/nosignificantdifference/>
- Saba, F. (1999). Is distance education comparable to "traditional education"?. Retrieved August 10, 2001 from the World Wide Web: <http://www.distance-educator.com/der/comparable.html>
- Schulman, A.H., Sims, R.L. (1999). Learning in an online format versus an in-class format: An experimental study. T.H.E. Journal, 26(11), 54-56.
- Sener, J., Stover, M. (2000). Integrating ALN into an independent study distance education program: NVCC case studies. Journal of Asynchronous Learning Networks, 4(2).
- Smeaton, A., Keogh, G. (1999). An analysis of the use of virtual delivery of undergraduate lectures. Computers and Education, 32, 83-94.
- Wade, W. (1999). Assessment in distance learning. T.H.E. Journal, 27(3), 94-100.