THE EFFECTS OF DISTANCE EDUCATION ON K–12 STUDENT OUTCOMES: A META-ANALYSIS WITH RECOMMENDATIONS FOR POLICY

Cathy Cavanaugh
University of North Florida

Kathy Jo Gillan
Florida Virtual School

Jeff Kromrey
University of South Florida

Melinda Hess
University of South Florida

Robert Blomeyer
Learning Point Associates

Keywords: virtual school, distance education, meta-analysis

Abstract

This meta-analysis is a statistical review of 14 studies with 116 effect sizes from web-delivered K–12 distance education programs studied between 1999 and 2004. The analysis shows that distance education can have the same effect on measures of student academic achievement when compared to traditional instruction. The study-weighted mean effect size across all outcomes was -0.028. No factors were found to be related to significant positive or negative effects. The factors that were tested included academic content area, grade level of the students, role of the distance learning program, role of the instructor, length of the program, type of school, frequency of the distance learning experience, pacing of instruction, timing of instruction, instructor preparation and experience in distance education, and the setting of the students.

Online virtual schools may be ideally suited to meet the needs of stakeholders calling for school choice, high school reform, and workforce preparation in 21st century skills. The growth in the numbers of students learning online and the importance of online learning as a solution to educational challenges has increased the need to study more closely the factors that effect student learning in virtual schooling environments. This meta-analysis is a statistical review of web-delivered K–12 distance education programs conducted in order to determine how student learning in online programs compares to learning in classroom-based programs, to identify the specific factors that influence student learning, and to make recommendations for improving practice in K-12 distance education.

Distance Education in the K–12 Context

The many thousands of K–12 students who participate in online education programs are attracted to virtual schooling because it offers advantages over classroom-based programs. Among the benefits of distance education for school-age children are increased opportunity for enrollment in courses as education programs reach underserved regions, broader educational opportunity for students who are unable to attend traditional schools, access to resources and instructors not locally available, and increases in student-teacher communication. Students in Canada's virtual secondary schools who used a range of telecommunications technologies showed greater improvement that their conventional school counterparts in several learning and technology skills (Barker and Wendel 2001). Academic advantages over traditional classroom instruction were demonstrated by students who learned using broadcast television in Mexico's Telesecundaria program, who were "substantially more likely than other groups to pass a final 9th grade examination" administered by the state (Calderoni 1998, p. 6); and by students taking a chemistry course by satellite, when measured using the American Chemical Society's course exam (Dees 1994). Virtual school developers and instructors continue to refine their practice, and in so doing, they learn from reports of both successful and unsuccessful programs.

Virtual schooling, like classroom schooling, has had limited success in some situations. In an online environment, students may feel isolated, parents may have concerns about children's social development, students with language difficulties may experience a disadvantage in a text heavy online environment, and subjects requiring physical demonstrations of skill such as music, physical education, or foreign language may not be practical in a technology-mediated setting.

For example, Bond (2002) found that distance between tutor and learner in an online instrumental music program has negative effects on performance quality, student engagement, and development and refinement of skills and knowledge. While distance learning was viewed as beneficial for providing the opportunity for elementary school students to learn a foreign language, Conzemius and Sandrock (2003) report that "the optimal learning situation still involves the physical presence of a teacher" (p. 47). Virtual school students show less improvement than those in conventional schools in listening and speaking skills (Barker and Wendel 2001). Highly technical subjects such as mathematics and science can also be difficult to teach well online. The Alberta Online Consortium evaluated student performance on provincial end-of-year exams among virtual school students across the province, and found that virtual school student scores in mathematics at grades 3, 6, 9, and 12, and the sciences at grades 6 and 9 lagged significantly behind scores of non-virtual school students (Schollie 2001).

Given instruction of equal quality, groups of students learning online generally achieve at levels equal to their peers in classrooms (Kearsley 2000). Equality between online and face-to-face delivery systems has been well documented over decades for adult learners, and while much less research exists focusing on K–12 learners, the results tend to agree. Several studies report no significant differences between K–12 distance education and traditional education in academic achievement (Falck et al. 1997; Goc Karp and Woods 2003; Hinnant 1994; Jordan 2002; Kozma et al. 2000; Mills 2002; Ryan 1996), frequency of communication between students and teachers (Kozma et al), and attitude toward courses (McGreal 1994).

Although various forms of technology-enabled distance education for K-12 students have been in use for nearly a century, rapid change in technology and the educational context have resulted in a small body of research relevant to today's conditions that can serve to guide instructors, planners, or developers. The temptation may be to attempt to apply or adapt findings from studies of K–12 classroom learning or adult distance learning, but K–12 distance education is fundamentally unique.

Characteristics of K-12 Online Learners

A primary characteristic that sets K-12 distance learners apart from their adult counterparts is their level of autonomy, which is necessary for success as a distance learner (Keegan, 1996). By the time learners reach higher education, most have acquired a degree of autonomy in learning, but younger students need to be scaffolded as part of the distance education experience. Adult learners more closely approach expertise in the subjects they study and in knowing how to learn independently, due to their long experience with the concepts and with meta-cognition, whereas children are relative novices. Experts organize and interpret information very differently from novices, and these differences affect learners' abilities to remember and solve problems (Bransford, Brown, & Cocking, 1999), and their ability to learn independently. Expert learners have better developed metacognition, a characteristic that children develop gradually.

A second characteristic that differentiates K-12 distance learners from adult learners is their locus of control. An internal locus of control leads learners to persist in the educational endeavor (Rotter, 1989). Research has found that older children have more internal locus of control than younger children (Gershaw, 1989), reinforcing the need for careful design and teaching of distance education at K–12 levels. Effective online programs for young learners include frequent teacher contact with students and parents, lessons divided into short segments, mastery sequences so student progress can grow in stages, and rewards for learning.

The success of K-12 learners in the online environment depends on whether course designers, providers, and instructors account for the needs of the learners. This meta-analysis attempts to collect data about K-12 online learning programs in part to determine the extent to which children's needs are accommodated in the programs.

Purpose of the Study

The purpose of this meta-analysis is to provide a quantitative synthesis of the research literature of web-based K–12 distance education from 1999 to the present, across content areas, grade levels, and outcome measures. The first goal was to determine the effects of distance education on K–12 student outcomes, specifically academic achievement. The second goal was to identify the effects on K-12 student outcomes of the features of distance education, including content area, duration of use, frequency of use, grade level of students, role of the instructor, type of school, timing of interactions, and pacing of the learning. The third goal was to make policy recommendations to K-12 distance education developers, managers, and instructors based on these results that will lead to the improvement practice.

Because all of the studies included in this synthesis drew data from school-based classes, this review can provide valuable insight into the practical effectiveness of K–12 distance education. Each study in this synthesis was conducted as a quasi-experimental study in a functioning virtual course, rather than as a controlled experiment. Controlled experimental research may offer findings of theoretical interest but may not be generalizable to complex learning settings such as virtual schools or classes. The uncontrollable cultural and social variables naturally present in a school or class, whether online or on-ground, make a statistical synthesis a more exact test of the strength of K–12 distance education. The effects of virtual learning would have to be strong and consistent to be measurable across a range of natural milieus.

Method

Given sufficient quantity and detail in the data, meta-analysis is capable of not only comparing the effectiveness of distance education programs to classroom-based programs, but it can compare features of various distance education programs to learn what works. The stages of the meta-analysis were identification and retrieval of applicable studies, coding of study features and findings, and data analysis. For the purposes of this meta-analysis, studies were included in the analysis if they met the following criteria for inclusion. Each study must:

- Be available as a journal article, dissertation or report in English between 1999 and 2004.
- Compare K–12 students in a distance education group to a non-distance education group, or compare the distance education group before and after distance education.
- Use web-based telecommunications, such that at least 50 percent of the students' participation in the course or program occurred at a physical distance from the instructor. Students were not necessarily physically separated from each other.
- Be a quantitative, experimental, or quasi-experimental study for which effect size could be computed, the outcome measures were the same or comparable, and the N was 2 or greater.
- Use student academic achievement, motivation, attitude, retention, or conduct as outcome variables, using either pre-test/pot-test design, or post-test only.

Location and Selection of Studies

Numerous databases, journals, websites, and bibliographic resources were searched for studies that met the established inclusion criteria. In each case, search terms included: cybercharter, cyberschool, distance education, distance learning, elearning, mlearning, online school, open learning, open school, schoolnet, telelearning, virtual charter, and virtual school. The web sites of several distance education organizations and over 200 virtual schools were browsed for studies, and the director of each virtual school was contacted at the email address listed on the school's website to request studies. The department of education website for each state was browsed for report cards for state virtual charter schools. Fourteen studies (see Table 1) were found to meet all criteria for inclusion. The studies that were examined and excluded were descriptive reports, did not meet the definition of distance education, lacked control or comparison group data, or did not provide data sufficient to compute effect size.

Limitations of the Review

For literature on K–12 distance education to be meaningfully synthesized, the inclusion criteria had to be narrowly specified. Measures of performance present in the literature do not draw a complete picture of the full range of effects that students experience as a result of participation in distance education. Qualitative studies, strict experimental studies, narrative reports, and other designs offer information not acquired in this analysis. Although the inclusion criteria were designed to allow a wide range of studies to be analyzed so that a comprehensive knowledge of K–12 distance education would result, a small number of studies was analyzed. The results should be interpreted with caution. Because the inclusion criteria in this synthesis were intentionally broad in order to allow greater generalizability, the quality of the studies is mixed. For that reason, study quality features were investigated as part of the meta-analysis (Greenland, 1994).

Coding of Study Features

Coding of study features allows the meta-analyst to unravel different study factors related to variations in the phenomenon from factors related to method (Lipsey and Wilson 2001). The coding used in this analysis was identified from research on K–12 distance education and from variables typically coded in contemporary meta-analyses in education. Each study was coded independently by two researchers according to the established coding procedure. The initial inter-rater agreement across all coded variables was 85 percent. Discrepancies between researchers were discussed and resolved.

The dependent variable in this synthesis was student outcome measured by instruments appropriate to the individual study given at the end of the distance education period which varied from a few weeks to an entire academic year. The measures included district, state, or national examinations, as well as teacher or researcher designed tests of academic performance. The studies were coded on 45 factors, categorized into five groups: identification of studies, distance education features, instructor/program features, study quality features, and sources of invalidity. The coding for distance learning and instructor features is included as an appendix to this article. The reliability coefficients of the measures were included in the coding. Of particular interest were the variables associated with distance education features (e.g. duration of the experience, role of the distance learning, role of the instructor, timing of the interactions) and instructor/program features (e.g. amount of teacher preparation for distance teaching, setting of the students). In many cases, however, the literature failed to report the detail needed to make meaningful comparisons on these factors. This lack of detail needed to make recommendations about practice led to the development of the policy implications included in the Discussion section of this article. The levels of each variable were compared by computing average effect sizes for each level, but examination of interactions among the different variables was not practical due to the small number of effect sizes available.

Calculation of Effect Sizes

The effect sizes estimated for each study outcome were computed using Cohen's *d*, defined in this meta-analysis as the difference between the non-distance learning group and the distance learning posttest mean scores divided by the pooled standard deviation. A correction factor for small sample bias in effect size estimation (Hedges, Shymansky, and Woodworth 1989) was used in cases in which sample sizes were small. The unit of analysis was the study outcome. For studies in which more than one independent group of students was evaluated, independent effect sizes were estimated for each group, were weighted to avoid study bias, and were included in the aggregated effect size estimate.

Table 1. Selected study features and effect sizes for 14 studies of web-based K–12 distance education

Author, year	*Grade level*	*Subject area*	*N*	*Weighted mean effect size (d)*	*95% CI for d (upper/lower)*
Alberta Consortium 2001*	3, 6, 9, 12	English, Mathematics, Science, Social Studies	13-397	-0.028	0.141/-0.197
Alaska Department of Education and Early Development 2003*	4-7, 9-12	Reading, Writing, Mathematics	7-67	-0.005	0.303/-0.313
Colorado Department of Education 2003a*	3-6	Reading, Writing, Mathematics	33-45	-0.028	0.261/-0.276
Colorado Department of Education 2003b*	7-8	Reading, Writing, Mathematics	9-55	-0.029	0.199/-0.258
Colorado Department of Education 2003c*	3-6	Reading, Writing, Mathematics	14-23	-0.013	0.440/-0.466
Colorado Department of Education 2003d*	7-8	Reading, Writing, Mathematics	10-21	-0.013	0.449/-0.475
Goc Karp & Woods 2003*	9-12	Physical education	19	-0.253	0.357/-0.863
Indiana Department of Education, 2004*	3, 6	Reading, mathematics	17-18	0.001	0.470/-0.468
Minnesota Department of Education 2003*	5	Reading, mathematics	26	0.014	0.398/-0.371
Mock 2000*	12	Science	7	-0.472	0.472/-1.416
Stevens 1999*	12	Science	21-33	-0.029	0.497/-0.556
Washington Office of the Superintendent of Public Instruction 2003*	7	Reading, Writing, Mathematics, Listening	12-15	0.002	0.540/-0.537
Wisconsin Department of Public Instruction 2003	3	Reading	57	-0.016	0.243/-0.276
Texas Education Agency 2003*	9-11	English, Mathematics, Science, Social Studies	15-21	-0.014	0.445/-0.474

* Indicates reports with more than one result. The range for N indicates the sample sizes for the various results in the report.

Statistical Analysis of Effect Sizes

Prior to the analysis of effect sizes for the set of studies identified in the meta-analysis, a funnel plot was constructed for the evaluation of the potential for publication bias in the sample (Light and Pillemer 1984). This bivariate plot of observed effect sizes and effect size standard errors suggested that the mean effect size remained consistent across the magnitude of the standard error, thus indicating no apparent publication bias. However, methods for detecting publication bias have been shown to have little statistical power for confirming bias in small samples of studies (e.g., Kromrey and Rendina-Gobioff in press), and the potential for bias in the sample of available studies must be acknowledged.

The test for heterogeneity (Q), based on Hedges and Olkin (1985), was used to determine whether the effect sizes of the studies were homogenously distributed, in other words, to learn whether the distribution of effect sizes around their mean was what would be expected from sampling error alone (Lipsey and Wilson 2001). The Q value for the weighted effect sizes was 1.485, and was considered to be homogeneous, indicating that the variance observed was likely to be due to sampling error. Therefore, the fixed-effects model was used to estimate variance (Kromrey and Hogarty 2002). Despite the failure to reject the null hypothesis of the Q test of homogeneity, the known low statistical power of this test (Harwell 1997; Hedges and Vevea 1998; Hogarty and Kromrey 1999) suggests the need for direct, focused tests of moderator variables. Therefore, study feature analyses were performed to determine the extent to which student outcomes were moderated by the study variables. *Statistical Analysis System* (SAS) software was used for the analyses. In order to control for the potential bias that may accrue in the analysis of dependent effect sizes discussed earlier, the aggregated effect sizes were weighted by the inverse of the number of effect sizes extracted from each study (Becker, 2000).

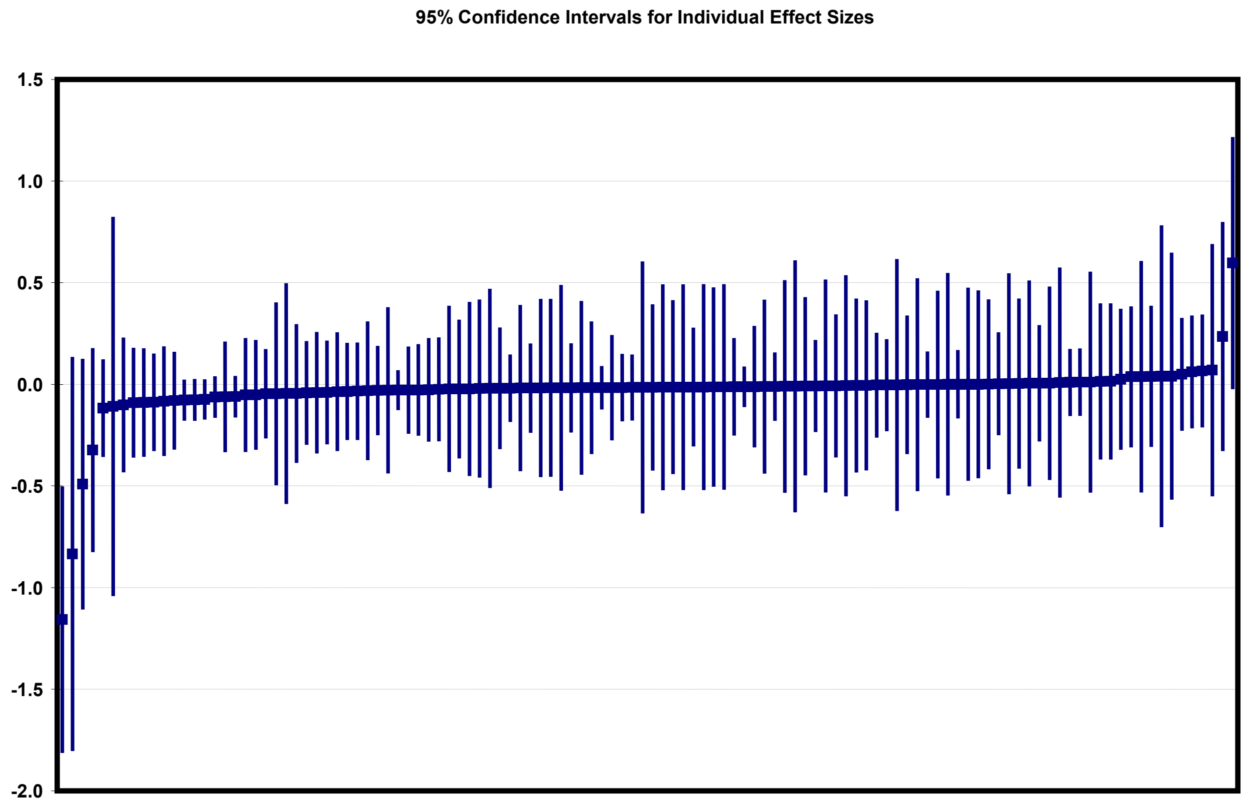
Results

Overall Effects of K–12 Distance Education

The 14 studies included in the analysis yielded 116 effect sizes drawn from a combined sample of 7561 students whose performance as a result of participation in a distance education program was compared to control groups in which students did not participate in distance education. The analysis resulted in an overall weighted effect size not significantly different from zero, a result that is consistent with the results of recent meta-analyses of distance education which tend to show that distance education is as effective as classroom instruction. The weighted mean effect size across all results was -0.028, with a standard error of 0.045 and a 95 percent confidence interval from -0.116 to 0.060. The average unweighted Cohen's *d* was -0.034, and the median effect size was -0.015. The effect sizes varied considerably among the studies. Unweighted effect sizes ranged from -1.158 to 0.597, with a standard deviation of 0.157, indicating that some applications of distance education appeared to be much better than classroom instruction and others were much worse.

The 95 percent confidence intervals also show wide variability in their size, as displayed in Figure 1. Only one confidence interval did not encompass zero, and all but three effect sizes fell between 0.5 and -0.5. Each of the fourteen studies and all except one of the 116 outcomes within the studies had individual effect sizes that did not differ significantly from zero, indicating that in almost every comparison, students in distance education programs performed as well as students in classroom-based programs.

Figure 1. 95 percent confidence intervals for individual effect sizes of 116 outcomes



Publication and Methodological Variables

Twenty variables were coded to discover whether publication or methodological variables accounted for variation in effect sizes. The publication features included the year of publication, the type of publication, and the region of publication. The methodological variables related to the testing sequence in the study, the type of achievement measure used in achievement studies, pretest equivalency measures, study design, statistical power, and control for 12 potential sources of invalidity. None of the variable comparisons resulted in effect sizes significantly different from zero.

Distance Education Variables

Eleven variables were used to identify the features of the distance education experience that may play a role in student performance. They were duration of the program, frequency of use of distance learning, instructional role of the program, number of distance learning sessions, duration of distance learning sessions, pacing of the instruction, role of the instructor, timing of the interactions, type of interactions, amount of teacher preparation for distance instruction, and amount of teacher experience in distance instruction. All levels of each distance education variable had effect sizes not significantly different from zero.

Instructional and Program Variables

The five variables that indicated the extent to which instructional and program factors played a role in student outcomes were grade level, school type, content area, the qualifications of the teacher in the teaching field, and the setting of the students. All instructional and program factors had effect sizes that were effectively zero.

Discussion

The questions of the effectiveness of distance education for K–12 student performance, and of the factors influencing its effectiveness were addressed using fixed-effects effect size estimation. The number of studies was small, and many studies did not report detailed information, so the results should be viewed as indications of tendencies rather than prescriptions for practice. Prior to this point, the field has relied on small individual studies, syntheses that included outdated analog technology, and syntheses that included adult learners.

The consistency of the effects shown in the studies analyzed in this review suggest that as distance education is currently practiced, educators and other stakeholders can reasonably expect learning in a well-designed distance education environment to be equivalent to learning in a well-designed classroom environment. How will K–12 distance education realize greater potential and maximize its effectiveness? How will designers and managers of K–12 distance education programs make better decisions in order to design and deliver a more effective program? The answers lie in changes in the ways policymakers and researchers do their work in this complex context. In order for distance education to be evaluated, data must be collected and reported in detail. The lack of detail in the research to date hinders thorough investigation of the factors influencing practice, and limits what can be learned for the improvement of practice. A coordinated research and reporting effort is needed in order to improve the cycle of conducting research on practice and applying research to improve practice.

Recommendations for K–12 Online Learning Policy and Practice

Optimally, the research on K–12 distance education would recommend specific practices that would lead to results that exceed those in conventional education settings. The barriers that prevent such recommendations include:

- a limit on the educational expertise focused on distance education as an area of study.

- a rather short-sighted view of the purposes of distance education, a lack of consensus about the goals of distance education, and an accompanying lack of evaluation directed at assessing progress toward those goals.
- a failure to take into account the complexity of systems in which distance education operates. Complexity is difficult to quantify, but virtual schooling evaluation and research can begin to track a greater range of influences, leading to a more thorough understanding of its effects.
- a paucity of research and reporting that includes details sufficient for quantitative synthesis.

For distance education to add a prospective agenda to the archive of valuable retrospective study that currently guides the field, five major action recommendation must be addressed by online learning practitioners, online learning district-level leadership, and Federal and State educational policy makers:

1. First, the broader educational community needs to become better informed about K–12 online learning and distance education, to foster better communication among the widest range of experts and practitioners who have the potential to contribute to advances in the field.

2. Second, the community of distance education policy makers, researchers, and practitioners should develop and articulate a long-range view of the intended and expected benefits of distance education and become advocates for suitably long-term studies of its effects.

3. Third, because education occurs in a dynamic context, and the rapid change in the technology used in distance education adds to the complexity, evaluation of distance education programs needs to account for more of this complexity than has so far been the practice. A common "codebook" or heuristic descriptive system should be created and refined to ensure that outcomes from distance and online learning programs can be accurately compared to other online and distance programs and to face-to-face instruction.

4. Standards are needed for reporting the academic and programmatic outcomes of distance education programs. Many K–12 distance education program directors collect admirable amounts of data, and conduct in-house analyses, but until there are standards set to guide the reporting of data, educational research will remain limited to examining results from only a small, enlightened subset of these programs.

5. The actions recommended require coordination and leadership. Leadership should begin at the national level and include the United States Department of Education, the North American Council on Online Learning and International Society for Technology in Education, in organizing a national distance learning community of practice to work toward enacting these essential action recommendations.

Learning, progress, and data-driven decisions require the availability of relevant data. The K–12 distance education and online learning communities certainly have the infrastructure for sharing that information. What is needed now is an adequate and uniform system for describing academic and programmatic outcomes within and across a variety of programs and instructional delivery systems, and uniform metrics and standards that can support comparisons within and across the various delivery systems and instructional modalities.

With ubiquitous availability of good information on the performance of all K–12 educational programs and instructional systems, parents and practitioners, policymakers and national political leadership will be able to make the very best informed decisions about how to best educate and equip all our children for life and success during the ensuing twenty-first century.

Acknowledgement: This work was originally produced by the North Central Regional Educational Laboratory with funds from the Institute of Education Sciences (IES), U.S. Department of Education, under contract number ED-01-CO-0011.

References

References marked with an asterisk indicate studies used in the meta-analysis.

*Alaska Department of Education and Early Development. 2003. Spring 2003 benchmarks, Delta/Greely school district - Delta Charter Cyber School. Available online at http://www.eed.state.ak.us/tls/assessment/AsmtVer/SchoolAsmtVerSuptSearch.cfm

Barker, K., and T. Wendel. 2001. E-learning: studying Canada's virtual secondary schools. Kelowna, BC: Society for the Advancement of Excellence in Education. Available online at http://www.excellenceineducation.ca/pdfs/006.pdf

Bond, A. 2002. Learning music online: An accessible program for isolated students. Kensington Park, SA: Australian National Training Authority. Available online at http://www.ncver.edu.au/research/proj/nr1013.pdf

Bransford, J., Brown, A., & Cocking, R. (Eds.). (1999). How people learn. Washington, DC: National Academies Press.

Calderoni, J. 1998. Telesecundaria: Using TV to bring education to rural Mexico. Education and Technology Technical Notes Series: World Bank Human Development Network. Available online at http://wbln0018.worldbank.org/HDNet/HDdocs.nsf/C11FBFF6C1B77F9985256686006DC949/1635F1703FE053B385256754006D8C3F/$FILE/telesecundaria.pdf

* Colorado Department of Education. 2003a. Branson alternative school, grades 1–6 school accountability report. Denver, CO: Author. Available online at http://reportcard.cde.state.co.us/reportcard/pdf/2003_1750_0948_E.pdf

* Colorado Department of Education. 2003b. Branson alternative school, grades 7–8 school accountability report. Denver, CO: Author. Available online at http://reportcard.cde.state.co.us/reportcard/pdf/2003_1750_0948_M.pdf

* Colorado Department of Education. 2003c. Connections academy, grades 1–6 school accountability report. Denver, CO: Author. Available online at http://reportcard.cde.state.co.us/reportcard/pdf/2003_0880_1887_E.pdf

* Colorado Department of Education. 2003c. Connections academy, grades 7–8 school accountability report. Denver, CO: Author. Available online at http://reportcard.cde.state.co.us/reportcard/pdf/2003_0880_1887_M.pdf

Conzemius, A., and P. Sandrock. 2003. Developing world language programs in elementary grades. Madison, WI: Wisconsin State Department of Public Instruction. ERIC, ED 480156.

Dees, S. 1994. An investigation of distance education versus traditional course delivery using comparisons of student achievement scores in advanced placement chemistry and perceptions of teachers and students about their delivery system (satellite course). Unpublished dissertation, Northern Illinois University, DeKalb.

Falck, A-K, J. Husu, T. Kronlund, H. Kynaslahti, J. Salminen, and M. Salonen. 1997. Testing virtual classroom in the school context. Distance Education. 18 (2): 213.

Gershaw, D. (1989). Line on life: locus of control. Retrieved May 29, 2004, from http://www3.azwestern.edu/psy/dgershaw/lol/ControlLocus.html

Goc Karp, G., and M. Woods. 2003. Wellness NutriFit online learning in physical education for high school students. Journal of Interactive Online Learning, 2 (2). Available online at http://www.ncolr.org/jiol/archives/2003/fall/03/index.html

Greenland, S. 1994. Invited commentary: A critical look at some popular meta-analytic methods. *American Journal of Epidemiology, 140*, 290-296.

Harwell, M. 1997. An empirical study of Hedges' Homogeneity Test. *Psychological Methods*, 2, 219-231.

Hedges, L., J. Shymansky, and G. Woodworth. 1989. Modern methods of meta-analysis. Washington, DC: National Science Teachers Association.

Hedges, L. V., & Vevea, J. L. 1998. Fixed- and random-effects models in meta-analysis. Psychological Methods 3, 486-504.

Hinnant, E. 1994. Distance learning using digital fiber optics: a study of student achievement and student perception of delivery system quality. Unpublished dissertation, Mississippi State University, Starkeville.

Hogarty, K. Y. & Kromrey, J. D. 1999. Traditional and robust effect size estimates: Power and Type I error control in meta-analytic tests of homogeneity. American Statistical

Association Proceedings of the Section on Government Statistics and Section on Social Statistics, 426-431.

* Indiana Department of Education. 2004. School snapshot: Irvington community school. Indianapolis, IN: Author. Available online at http://mustang.doe.state.in.us/SEARCH/snapshot.cfm?schl=1537

Jordan, A. 2002. An investigation into the effects of online teaching and learning on achievement outcomes at the secondary level. Unpublished dissertation, Fayetteville State University, Fayetteville, NC.

Kearsley, G. 2000. Online education. Belmont, CA: Wadsworth.

Keegan, D. (1996). Foundations of distance education. London: Routledge.

Kozma, R., A. Zucker, C. Espinoza, R. McGhee, L. Yarnall, D. Zalles, and A. Lewis. 2000. The online course experience: evaluation of the virtual high school's third year of implementation, 1999-2000. Arlington, VA: SRI. Available online at http://www.govhs.org/Images/SRIEvals/$file/SRIAnnualReport2000.pdf

Kromrey, J., and K. Hogarty. 2002. Estimates of variance components in random effects meta-analysis: sensitivity to violations of normality and variance homogeneity. American Statistical Association Joint Statistical Meeting, 1963–1968.

Kromrey, J. D. and Rendina-Gobioff, G. (in press). On knowing what we don't know: An empirical comparison of methods to detect publication bias in meta-analysis. Educational and Psychological Measurement.

Light, R. J. and Pillemer, D. B. 1984. Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press.

Lipsey, B., and D. Wilson. 2001. Practical meta-analysis. Thousand Oaks, CA: SAGE Publications, Inc.

McGreal, R. 1994. Comparison of the attitudes of learners taking audiographic teleconferencing courses in secondary schools in northern Ontario. Interpersonal Computing and Technology, 2 (4), 11–23. Available online at http://www.helsinki.fi/science/optek/1994/n4/mcgreal.txt

Means, B., and G. Haertel. 2004. A blueprint for a national research agenda to evaluate educational technology. In Using technology evaluation to enhance student learning, ed. B. Means, and G. Haertel. New York: Teachers College Press.

Mills, S. 2002. School isn't a place anymore: an evaluation of virtual Greenbush online courses for high school students. University of Kansas: Schiefelbusch Institute for Life Span

Studies. Available online at http://media.lsi.ku.edu/research/vgeval/vgevaluationreport.pdf

* Minnesota Department of Education. 2003. 2003 report card: cyber village academy. Minneapolis, MN: Author. Available online at http://education.state.mn.us/ReportCard/2003/RCF402507010.pdf

* Mock, R. 2000. Comparison of online coursework to traditional instruction. Unpublished thesis, Michigan State University, East Lansing. Available online at http://hobbes.lite.msu.edu/~robmock/masters/abstract.htm

Rotter, J. B. (1989). Internal versus external control of reinforcement: A case history of a variable. American Psychologist, 45, 489–493.

Ryan, W. 1996. The effectiveness of traditional vs. audiographics delivery in senior high advanced mathematics course. Journal of Distance Education/Revue de l'enseignement à distance (11)2. Available online at http://www.cade-aced.ca/en_pub.php

Sabelli, N. 2004. Policy, planning, and the evaluation of learning technology. In Using technology evaluation to enhance student learning, ed. B. Means, and G. Haertel. New York: Teachers College Press.

* Schollie, B. 2001. Student achievement and performance levels in online education research study. Edmonton, Alberta: Alberta Online Consortium. Available online at http://www.albertaonline.ab.ca/pdfs/AOCresearch_full_report.pdf

* Stevens, K. 1999. Two Canadian approaches to teaching biology, chemistry, mathematics and physics to senior high school students in virtual classes. Australasian Science Education Research Association. Available online at http://www.tellearn.mun.ca/pubs/virtual.html

* Texas Education Agency. 2003. 2002–2003 campus performance: Southwest Virtual Preparatory School. Austin, TX: Author. Available online at http://www.tea.state.tx.us/perfreport/aeis/2003/campus.srch.html

* Washington Office of the Superintendent of Public Instruction. 2003. Internet academy 2002–2003 WASL results. Olympia, WA: Author. Available online at http://reportcard.ospi.k12.wa.us/Reports/summary.aspx?schoolId=1164&reportLevel=School

* Wisconsin Department of Public Instruction. 2003. Successful school guide: Wisconsin Virtual Academy. Madison, WI: Author. Available online at http://data.dpi.state.wi.us/data/questions.asp?fullkey=011945040100&DN=Northern+Ozaukee&SN=Wisconsin+Virtual+Academy&TYPECODE=6&CTY=45&ORGLEVEL=SC

Appendix 1. Meta-Analysis Coding for Distance Learning and Instructor Features

A. Identification of studies

Grade level of students. Unspecified=00, grades 1–12 use 01 to 12, Mixed primary (K–2) =13, Mixed intermediate (3–5) =14, Mixed middle (6–8) =15, Mixed high (9–12) =16, K–12=17, other=18.

School type. Unspecified=0, Public district sponsored=1, Public state sponsored=2, Private=3, Other=4, Charter=5, Combination=6.

B. Distance learning features

Duration of distance learning experience. Less than one semester=1, One semester=2, More than one semester=3.

Frequency of distance learning experience. Unspecified=0, From 5 to 7 days per week=1, From 1 to 4 days per week=2, From 1 to 3 days per month=3, Less than monthly=4.

Instructional role of distance learning. Unspecified=0, Full-time educational program=1, Courses to supplement an educational program or partial educational program=2, Supplement to a specific course=3.

Number of distance learning sessions. Unspecified=0, List number of sessions.

Duration of distance learning sessions. Unspecified=0, List average minutes per session.

Pacing of distance learning instruction. Unspecified=0, Completely self-paced=1, Student sets pace within instructor-determined parameters=2, Pacing completely specified by program or instructor=3.

Instructor role. Unspecified=0, Fully moderated=1, Unmoderated=2, Combination=3, Other=4.

Timing of interactions. Unspecified=0, Synchronous=1, Asynchronous=2, Combination=3, Other =4.

Type of interactions. Unspecified=0, Student—content=1, Student—instructor=2, Student—student=3, Student—others=4, Combination=5, Other=6.

C. Instructor/program features

Amount of teacher preparation in distance learning. Unspecified=0, List hours of preparation.

Amount of teacher experience in distance learning. Unspecified=0, List years of experience.

Qualifications of teacher in the teaching field. Unspecified=0, Certified in content area=1, Certified but teaching out of field=2, Alternative or provisional certification=3, Uncertified=4, Other=5.

Setting of students during distance learning. Unspecified=0, Home=1, School=2, Other=3, Combination=4.