The Effects of Distance Education on K–12 Student Outcomes: A Meta-Analysis

October 2004

LEARNING POINT Associates

The Effects of Distance Education on K–12 Student Outcomes: A Meta-Analysis

October 2004

Cathy Cavanaugh
University of North Florida
Kathy Jo Gillan
Duval County Public Schools
Jeff Kromrey
University of South Florida
Melinda Hess
University of South Florida
Robert Blomeyer
North Central Regional Educational Laboratory

LEARNING POINT Associates

1120 East Diehl Road, Suite 200
Naperville, Illinois 60563-1486
(800) 356-2735 • (630) 649-6500
www.learningpt.org

This work was originally produced in whole or in part 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. The content does not necessarily reflect the position or policy of IES or the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

Contents

Abstract

The community of K–12 education has seen explosive growth over the last decade in distance learning programs, defined as learning experiences in which students and instructors are separated by space and/or time. While elementary and secondary students have learned through the use of electronic distance learning systems since the 1930s, the development of online distance learning schools is a relatively new phenomenon. 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 affect student learning in virtual schooling environments. This meta-analysis is a statistical review of 116 effect sizes from 14 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 with a 95 percent confidence interval from 0.060 to -0.116, indicating no significant difference in performance between students who participated in online programs and those who were taught in face-to-face classrooms. 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.

Introduction

The community of K–12 education has seen explosive growth over the last decade in distance learning programs, defined as learning experiences in which students and instructors are separated by space and/or time. While elementary and secondary students have learned through the use of electronic distance learning systems since the 1930s, the development of online distance learning schools is a relatively new phenomenon. 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.

Beginning in the 1930s, radio was used simultaneously to bring courses to school students and to help teachers learn progressive Deweyan methods of teaching (Bianchi, 2002), in what might have been among the earliest professional development school models. From that point on, television, audio and videoconferencing, the Internet, and other technologies have been adapted for the needs of young learners. This meta-analysis is a statistical review of web-delivered K–12 distance education programs between 1999 and 2004 conducted in order to determine how student learning in online programs compares to learning in classroom-based programs, and to identify the specific factors that influence student learning.

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 increases in enrollment or time in school 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 virtual schools showed greater improvement that their conventional school counterparts in critical thinking, researching, using computers, learning independently, problem-solving, creative thinking, decision-making, and time management (Barker & Wendel, 2001). Academic advantages over traditional classroom instruction were demonstrated by students 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); by students taking a chemistry by satellite course (Dees, 1994); and by students learning reading and math via interactive radio instruction (Yasin & Luberisse, 1998). 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 & Wendel, 2001). Highly technical subjects such as mathematics and science have also proven to be difficult to teach well online. The Alberta Online Consortium evaluated student performance on 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 nonvirtual 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 the 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. "Evidence to date convincingly demonstrates that , when used appropriately, electronically delivered education—'e-learning'—can improve how students learn, can improve what students learn, and can deliver high-quality learning opportunities to all children" (National Association of State Boards of Education, 2001, p. 4). Many studies report no significant differences between K–12 distance education and traditional education in academic achievement (Falck et al, 1997; Goc Karp & 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 pre-college 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 for Success

A primary characteristic that sets successful distance learners apart from their classroom-based counterparts is their autonomy (Keegan, 1996) and greater student responsibility (Wedemeyer, 1981). By the time they reach higher education, most adults have acquired a degree of autonomy in learning, but younger students need to be scaffolded as part of the distance education experience. Virtual school teachers must be adept at helping children acquire the skills of autonomous learning, including self-regulation. Adult learners more closely approach expertise in the subjects they study and in knowing how to learn, due to their long experience with the concepts and with meta-cognition, whereas children are relative novices. This distinction is important because 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 successful distance learners from unsuccessful ones is an internal locus of control, leading them 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. Younger students will need more supervision, fewer and simpler instructions, and a more extensive reinforcement system than older students. 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 such as multimedia praise and printable stickers or certificates.

Young students are different from adult learners in other ways. Piaget's stages of cognitive development, in particular preoperational (2 to 7 years), concrete operational (7 to 11 years), and formal operational (11 years to adulthood) outline the phases in development toward adulthood. The stages offer pedagogical guidance for delivering effective web based education, which should focus on the major accomplishments of learners in these stages. Each stage is characterized by the emergence of new abilities and ways of processing information (Slavin, 2003, p. 30), which necessitates specialized instructional approaches and attention to each child's development. Since adults have progressed through these stages of cognitive development, delivery of web based education at the adult level need not concentrate on methods that help the learner develop these cognitive skills. In contrast, web-based instruction for students in their formative years must include age appropriate developmental activities, building on the students' accomplishments in and through the cognitive stages. For example, an online mathematics or science lesson designed for students at the preoperational stage needs to use very concrete methods, such as instructing the student to develop concepts by manipulating and practicing with real-world objects. The concept can built upon for students in the concrete operational stage using multimedia drag-and-drop manipulations and representations, or realistic simulations. At the formal operational stage, students are capable of using symbols, language, and graphic organizers to continue to learn the concepts in more abstract ways.

Teaching and Learning Theory

Piaget helps us to understand that learning should be holistic, authentic, and realistic. Less emphasis should be placed on isolated skills aimed at teaching individual concepts. Students are more likely to learn skills while engaged in authentic, meaningful activities. Authentic activities are inherently interesting and meaningful to the student. Web-based technology offers a vast array of opportunities to help expand the conceptual and experiential background of the student (Bolton, 2002, p. 5).

Neo-Piagetian theorists have expanded on Piaget's model of cognitive development. Among others, Vygotsky proposed that historical and cultural context play significant roles in helping people think, communicate, and solve problems, proposing that cognitive development is strongly linked to input from others. Vygotsky's theory implies that cognitive development and the ability to use thought to control our own actions require first mastering cultural communication systems and then learning to use these systems to regulate our own thought process. He believed that learning takes place when children are working within their zone of proximal development. Tasks within the zone of proximal development are ones that children cannot yet do alone but could do with the assistance of more competent peers or adults (Slavin, 2003, p. 43–44). When working with children using web-based technology, teachers must offer students activities that make use of the web's powerful tools for collaborative learning, and are within their zone of proximal development. Online communities can provide a supportive context that makes new kinds of learning experiences possible (Bruckman, 1998, p. 84–85).

Constructivism, a widely used theory in distance education, is founded on the premises that by reflecting on our experiences and participating in social-dialogical process (Duffy & Cunnigham, 1996), we construct our understanding of the world we live in. Each of us generates our own "rules" and "mental models," which we use to make sense of our experiences. Learning, therefore, is simply the process of adjusting our mental models to accommodate new experiences (Brooks & Brooks, 1993). Children have not had the experiences that adults have had to help them construct understanding. Therefore, children construct an understanding of the world around them that lacks the rich experiences that adults have had. Scaffolding or mediated learning is important in helping children achieve these cognitive understandings (Slavin, 2003, p. 259), and are essential components of web-based learning experiences for children. Online learning environments, when designed to fully use the many tools of communication that are available, is often a more active, constructive, and cooperative experience than classroom learning. In addition, technologies that are easily employed in online environments, such as mind mapping tools and simulations, are effective means for helping students make meaning of abstract phenomena and strengthen their meta-cognitive abilities (Duffy & Jonassen, 1992).

Purpose of the Study

With the emphasis on scientifically-based research and the call for evidence-based program decisions in the federal No Child Left Behind Act of 2001, scientific evidence is needed to guide the growing numbers of online school developers and educators. Many studies of K–12 distance education have been published, but a small proportion of them are controlled, systematic, empirical comparisons that fit the definition of "scientific," as it is defined by the U.S Department of Education and described at the What Works Clearinghouse website, http://www.w-w-c.org/. This study is an effort to search for and collect the studies that fit the definition of scientific research on K–12 distance education programs, and to draw conclusions about the effectiveness of distance education for K–12 students based on the synthesized findings of the studies.

Meta-analysis is an established technique for synthesizing research findings to enable both a broader basis for understanding a phenomenon and a parsing of influences on the phenomenon. Several recent meta-analyses related to distance education have been published in recent years (see Table 1).

Table 1
Summary of recent meta-analyses in distance education

Author(s), Date	**Focus**	**N of studies**	**Effect Size**
Machtmes & Asher, 2000.	Adult telecourses	30	-0.0093
Cavanaugh, 2001.	Academic achievement of K–12 students	19	+0.015
Allen, Bourhis, Burrell, & Mabry, 2002.	Student satisfaction among adult learners	25	+0.031
Bernard, Abrami, Lou, Borokhovski, Wade, Wozney, Wallet, Fiset, & Huang, 2003.	Student achievement, attitude, retention	232	+0.0128
Shachar & Neumann, 2003.	Student achievement	86	+0.37
Ungerleider & Burns, 2003.	Networked and online learning	12 for achievement 4 for satisfaction	0 for achievement -0.509 for satisfaction

The Sachar and Neumann study was the only one to have found a moderate effect for distance education. Only one of the recent meta-analyses in distance education focused on K–12 learners, and it included web-based programs along with the analog conference and broadcast programs no longer in common use in today's virtual schools. The explosion in virtual schools, especially virtual charter schools in the United States, has necessitated a fresh look at the knowledge base. The need is for research that guides practitioners in refining practice so the most effective methods are used. 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. For example, synchronous programs can be compared to nonsynchronous programs. Meta-analysis is a tool that allows looking in detail at virtual schooling practice and results, and it can lead to better informed practice and improved results.

Several advantages can result from a synthesis of studies of the effectiveness of distance education programs for K–12 learners. Because all of the studies included in this review drew data from school-based classes, the review can provide valuable insight into the practical effectiveness of K–12 distance education. 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.

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 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.

From the literature, the meta-analysis seeks to answer the following questions:
1. Is distance education as effective, in terms of student achievement, as classroom-based instruction?
2. To what extent are student outcomes related to the features of a distance education system (duration of use, frequency of use, role of the instructor, timing of interactions, and pacing of the learning)?
3. To what extent are student outcomes related to features of the educational context (content area, school type, and grade level)?
4. To what extent are results related to study features (year, type of publication, various potential threats to validity)?

Meta-analysis, the use of statistical analysis to synthesize a body of literature, is appropriate for answering questions such as these because it allows comparison of different studies by computing an effect size for each study. Meta-analysis is used to estimate the size of a treatment's effect, and allows investigation into relationships among study features and outcomes (Bangert-Drowns, 2004). The inclusion of a study in a meta-analysis is limited by several factors, the most significant of which is the reporting of the information needed to compute effect size. Very often, reports released by virtual schools and other distance education programs do not include mean scores, comparison group scores, sample sizes, or standard deviations. Nonetheless, the meta-analytic technique is a way to identify effects or relationships in literature that may not be evident otherwise (Lipsey & Wilson, 2001).

Method

This quantitative synthesis is a meta-analysis of empirical studies published since 1999 that compared the effects of web-delivered distance education with classroom-based learning on K–12 student academic performance. Since 1999 the sophistication in the use of distance learning tools has improved, but the types of tools available to schools have remained approximately the same. The stages of the meta-analysis were identification and retrieval of applicable studies, coding of study features and findings, and data analysis. These stages are described below.

For the purposes of this meta-analysis, studies were included in the analysis if they met the following criteria for inclusion. The studies 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 nondistance 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.

- Be quantitative, experimental, or quasi-experimental studies 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.

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
- virtual school.

Electronic searches were systematically conducted in the following databases:

- *Dissertation Abstracts*
- *ERIC*
- *JSTOR*
- *Kluwer*
- *ProQuest Education*
- *PsychInfo*
- *Wilson Education.*

Web searches were performed using the Google, Teoma, Grokker, MetaCrawler, and AltaVista search sites.

Abstracts in the following distance education journals were examined:

- *American Journal of Distance Education*
- *Computers & Education*
- *Distance Education*
- *Journal of Distance Education*
- *Journal of Distance Learning*
- *Open Learning.*

Abstracts in the following educational technology journals were examined:

- *Association for the Advancement of Computing in Education* journals
- *Australasian Journal of Educational Technology*
- *British Journal of Educational Technology*
- *Canadian Journal of Educational Communication*
- *Canadian Journal of Learning and Technology*
- *Computers in the School*
- *Educational Technology & Society*
- *Educational Technology Research and Development*
- *Journal of Computer Mediated Communication*
- *Journal of Computing in Childhood Education*
- *Journal of Educational Computing Research*
- *Journal of Information Technology Education*
- *Journal of Interactive Media in Education*
- *Journal of Research on Technology in Education.*

Abstracts in *American Educational Research Journal* were examined, as were abstracts in the following electronic journals:

- *Australian Educational Computing*
- *Australian Journal of Educational Technology*
- *Electronic Journal for the Integration of Technology in Education*
- *International Journal of Educational Technology*
- *International Review of Research in Open and Distance Learning*
- *Journal of Asynchronous Learning Networks*
- *Journal of Interactive Online Learning*
- *Online Journal of Distance Education Administration*
- *TechKnowLogia: International Journal of Technologies for the Advancement of Knowledge and Learning*
- *Turkish Online Journal of Distance Education.*

In addition, abstracts were examined in the following conference proceedings:

- American Education Research Association
- Canadian Association for Distance Education
- EdMedia
- E-Learn/WebNet
- Society for Technology in Teacher Education.

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. The reference lists of the six recent meta-analyses of distance education shown in Table 1 were reviewed for potential studies.

Of the thousands of abstracts that were reviewed, 80 full-text articles, dissertations, or reports concerning DE and traditional instruction at K–12 level were obtained and evaluated for inclusion in the analysis. Independently, two researchers read all collected studies to determine eligibility for inclusion based on the stated criteria. Fourteen of the studies were found to meet all criteria for inclusion. Of the 66 studies that were examined and excluded, 28 percent were descriptive reports, 14 percent reported on uses of telecommunications or other educational technology that did not meet the definition of distance education, 25 percent reported results without control or comparison group data, and 33 percent included summary data only 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. This synthesis included studies with data on the performance of grades 3–12 students in web-based distance learning programs compared to students in classrooms. 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.

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 & 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. A trial conducted on a small sample of studies led to the addition of variables in the codebook that were not present in the initial set of variables. Each study was coded independently by two researchers according to the established coding procedure. The full codebook is included in Appendix A. The initial inter-rater agreement across all coded variables was 85 percent. Discrepancies between researchers were discussed and resolved. The entire dataset was reviewed for the presence of discrepancies and unexpected values.

Fourteen studies, with a total of 116 outcomes, had data sufficient to include in the analysis (see Table 2). 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 (see Appendix A). 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. 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 nondistance learning group and the distance learning posttest mean scores divided by the average standard deviation. A correction factor for small sample bias in effect size estimation (Hedges, Shymansky, & 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. A positive effect size, with a 95 percent confidence interval not encompassing zero, is an indication that the distance learning group outperformed the nondistance learning group.

Table 2
Selected study features and effect sizes for 14 studies of web-based K–12 distance education

Author, year	*Grade level*	*Subject area*	*School type*	*Outcome measure*	*Instructional role of the distance learning*	*Timing of interactions*	*N*	*Weighted mean effect size (d)*	*95% CI for d (upper/lower)*
Alberta Consortium 2001*	3, 6, 9, 12	English, mathematics, science, social studies	Mix of public and private	National test	Course	Asynchronous	13–397	-0.028	0.141/-0.197
Alaska Department of Education and Early Development 2003*	4–7, 9–12	Reading, writing, mathematics	State charter	State and national tests	Full program	Synchronous	7–67	-0.005	0.303/-0.313
Colorado Department of Education 2003a*	3–6	Reading, writing, mathematics	State charter	State test	Full program	Asynchronous	33–45	-0.028	0.261/-0.276
Colorado Department of Education 2003b*	7–8	Reading, writing, mathematics	State charter	State test	Full program	Asynchronous	9–55	-0.029	0.199/-0.258
Colorado Department of Education 2003c*	3–6	Reading, writing, mathematics	State charter	State test	Full program	Combination synchronous asynchronous	14–23	-0.013	0.440/-0.466
Colorado Department of Education 2003d*	7–8	Reading, writing, mathematics	State charter	State test	Full program	Combination synchronous asynchronous	10–21	-0.013	0.449/-0.475
Goc Karp &	9–12	Physical	Public	Class	Portion of	Asynchronous	19	-0.253	0.357/-

Woods 2003*		education		assignments	course				0.863
Indiana Department of Education, 2004*	3, 6	Reading, mathematics	State charter	State test	Full program	Unspecified	17–18	0.001	0.470/-0.468
Minnesota Department of Education 2003*	5	Reading, mathematics	State charter	State test	Full program	Unspecified	26	0.014	0.398/-0.371
Mock 2000*	12	Science	Public	Teacher made test	Portion of course	Asynchronous	7	-0.472	0.472/-1.416
Stevens 1999*	12	Science	Public	Teacher made test	Portion of course	Unspecified	21–33	-0.029	0.497/-0.556
Washington Office of the Superintendent of Public Instruction 2003*	7	Reading, Writing, mathematics, listening	State charter	State test	Full program	Asynchronous	12–15	0.002	0.540/-0.537
Wisconsin Department of Public Instruction 2003	3	Reading	State charter	State test	Full program	Asynchronous	57	-0.016	0.243/-0.276
Texas Education Agency 2003*	9–11	English, mathematics, science, social studies	State charter	State test	Full program	Combination	15–21	-0.014	0.445/-0.474

* indicates studies yielding multiple effect sizes

Statistical Analysis of Effect Sizes

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 & 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 & Hogarty, 2002). 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. Effect size comparisons were done for the variables: grade level, content area, duration and frequency of the distance learning experience, instructional role of the distance education, pacing of the instruction, role of the instructor, timing of the interactions, and types of interactions, as well as for various study quality and invalidity factors.

Results

Characteristics of the Studies

The 14 studies included in the analysis yielded 116 independent 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. Sixty one percent of the study results had sample sizes of less than 50, and 16 percent had sample sizes above 100. All but one of the studies included more than one comparison, and the average number of comparisons per study was 8, ranging from one to 38. Eighty six percent of the studies were organization reports, 7 percent were published articles, and 7 percent were dissertations. All of the studies were published between 1999 and 2004, with eleven published during 2003 and 2004, and three published from 1999 through 2001. Two studies were published in Canada, and the other twelve were published in the U.S.

A range of distance learning structures was examined in the literature. Half of the studies reported on programs that used asynchronous timing in instruction. Three studies documented a program that used a combination of synchronous and asynchronous instruction, one program was delivered synchronously, and the remaining programs did not report on instructional timing. Ten of the studies reported results of student participation in full year-long distance learning programs, one included data for distance learning courses, and three studies focused on portions of courses delivered at a distance for less than a semester. Thirteen studies included data from programs in which students participated approximately five days per week, and the other study did not indicate the frequency of student participation. The diversity of distance learning structures is an indication of the wide range of educational uses to which it is being applied: enhancement or extension to classroom instruction, school courses, and full-time educational programs.

The studies encompassed a variety of instructional features. The bulk of the results, 75 percent, occurred in the secondary grades, 6–12. The other results concern elementary age children, in grades 3–5. Results from seven academic content areas were reported. Thirty percent of the results came from tests of reading ability, followed by mathematics, which accounted for 26 percent of the results. Writing was the subject for 16 percent of the results, science was the topic of 14 percent, and social studies made up 9 percent of the results. Three percent of results came from physical education comparisons, and one percent from a test of listening. National tests were used to compare outcomes in one study, state tests were used in nine studies, teacher made tests were used in two studies, and one study reported data from both state and national tests.

Overall Effects of K–12 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 (see Table 1), 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. Figure 1 displays the full range of effect sizes calculated for the 116 results across the horizontal axis, and the number of results having each effect size on the vertical axis. The spike in the number of results around the zero effect size is an indicator of the tendency of the overall effect size. 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.

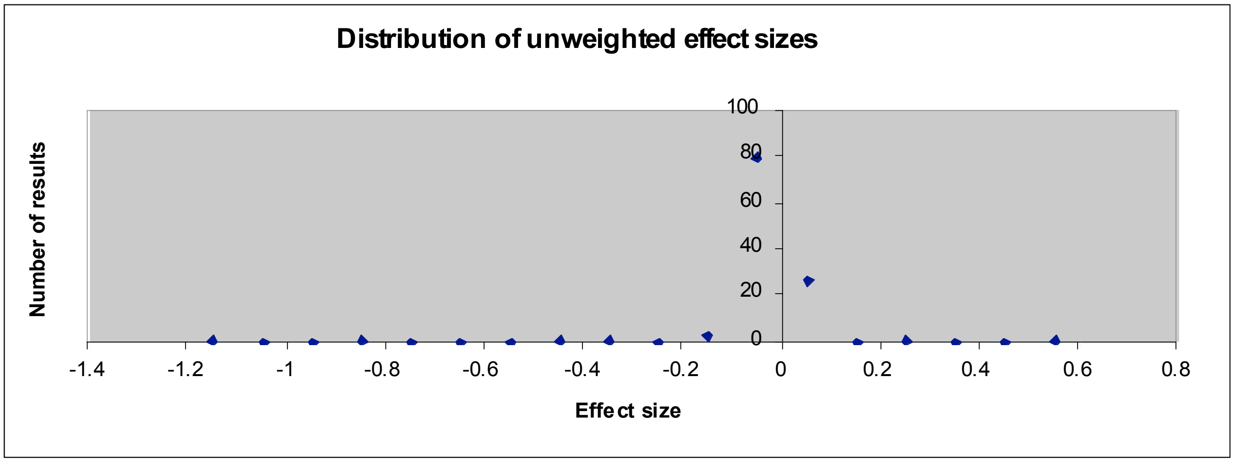

Figure 1. Distribution of unweighted effect sizes of 116 outcomes

The 95 percent confidence intervals also show wide variability in their size, as displayed in Figure 2. 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.

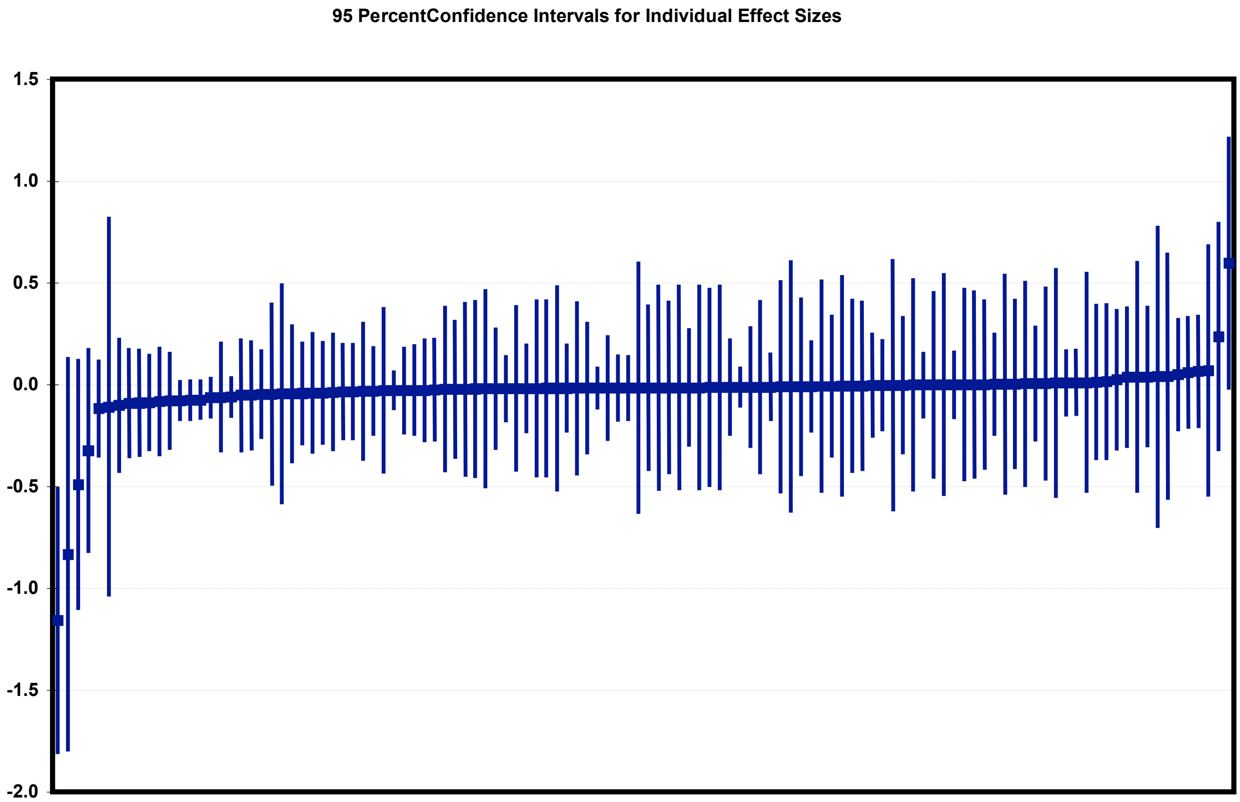

Figure 2. 95 percent confidence intervals for individual effect sizes of 116 outcomes.

Of the 45 factors coded in the study, the following 30 were examined to determine sources of significant variation in effect sizes. Ten of the remaining variables were used for identifying the studies or computing effect size, and the other five could not be compared because the studies did not include the data for coding the variables, or the variable was not a relevant factor in the studies. The variables that went uncoded due to the absence of data were the frequency of student participation in distance learning, the level of preparation of the teachers in distance education, and the amount of experience of the teachers in distance education. The variables that were not relevant factors for the studies were control for the effects of a second testing, and control for the effects of a pretest. Analysis of variance was not meaningful for some of the variables because of missing data in the studies, resulting in a high number of cases in which a value of "unspecified" was coded for the variable.

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 (see Table 2 and Figure 2).

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. Because of the individualized nature of distance education, only two of the studies indicated specific numbers and durations of distance learning sessions, and they were studies of limited partial-course experiences. Half of the studies did not indicate whether students or instructors set the pace within the distance learning timeframe, while three of the programs were completely self-paced, and four were designed for students to set their pace within parameters set by the instructor. In terms of the role of the instructor in teaching, one program was fully moderated, five were nonmoderated, four used a combination of moderated and nonmoderated activities, and four did not indicate the instructors' role. Ten programs used a combination of interactions among students, content, instructors, and others; one limited interactions to student-content; and three did not specify interaction types. No studies described the levels of instructor preparation or experience required of or possessed by the instructors. 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. Twelve of the studies indicated that the instructors were certified teachers, and the other two studies did not describe the credentials of the instructors. In five of the programs, students participated from home or a nonschool location, four programs are designed such that students completed some work from home and some in a school setting, in three programs, students did their distance learning work while at a school, and two programs did not specify the setting of the students. All instructional and program factors had effect sizes that were effectively zero.

Discussion

The literature reviewed in this meta-analysis includes results from 116 comparisons of grades 3–12 web-based distance education programs with classroom-based teaching, including data for 7561 students. 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 findings confirm those of other recent meta-analyses of distance education programs, and provide a needed update to the meta-analysis focused on K–12 students which was completed in 1998 just as the web-based systems were beginning to be studied in virtual schooling. The analysis showed that for the factors examined, distance learning did not outperform or underperform classroom instruction. 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. What has been learned from these results is that, based on the best research available on online K–12 distance education programs, such programs are effective for student learning. Prior to this point, the field has relied on small individual studies, syntheses that included outdated analog technology, and syntheses that included adult learners.

Implications for Research and Practice

Distance education as it has been implemented at the K–12 level over the past decade has improved over time according to several measures: providing access to education and choice in course offerings to increased numbers of students, offering education to a larger range of grade levels and ability levels, using more interactive and widely accessible technologies, and leading students to academic success on a wider range of achievement instruments. The effect of distance education on learning may be moderated by several factors, existing as it does in a very complex web of educational, technological, and social dynamics. Factors such as the design of the distance learning system, the demands of the content, the abilities and disabilities of the student, and the quality of the teacher are likely to be influential factors, as they are in conventional educational enterprises. 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 it 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. Such data collection begins with identification of goals. Policymakers and evaluators must enter into a partnership in which common goals are identified, an evaluation plan is acted on, and detailed reporting follows. Evaluation must be seen as a tool to support policy setting and decision making (Means & Haertel, 2004). It is no longer enough to ask whether distance education is effective, we need to understand why (Sabelli, 2004). We need to know how to make it more effective, what factors contribute most to effectiveness, and in what contexts the factors operate. Acquiring this knowledge requires consensus on a definition of effectiveness that goes beyond standardized tests, and a system for identifying and measuring factors that influence effectiveness. As Means and Haertel stress, "many studies of the effects of technology-supported innovations are hindered by a lack of measures of student learning commensurate with the initiative's goals" (p. 99).

One factor warranting special consideration in assessing the effectiveness of virtual schooling is teacher quality. In classrooms, teacher effectiveness is a strong determiner of differences in student learning, far outweighing differences in class size and heterogeneity (Darling-Hammond, 2000). Based on the similarities in student outcomes between distance and classroom learning, there is every reason to expect that teacher preparation is critical in distance education. However, there has been very little formal preparation available addressing the unique nature of online instruction and very little time for teachers to develop their expertise as online instructors. As professional development becomes more common and expertise grows, student success is likely to grow as well.

As second factor that is growing in importance in K–12 distance education is the emergence of virtual charter schools. By 2002, there were about 2000 charter schools nationwide, and the No Child Left Behind Act allows public schools that "chronically fail" to make adequate yearly progress to be restructured as charter schools (Nelson, Rosenburg, & Van Meter, 2004, p. 1). According to state department of education websites, there are now almost 100 virtual charter schools operating. This synthesis includes data from ten virtual charter schools, all of which performed at levels equivalent to nonvirtual public schools in their states. In contrast, the 20004 report on charter school achievement on the National Assessment of Education Progress (Nelson et al) provides evidence that charter schools overall are underperforming when compared to noncharter public schools. Charter school students had significantly lower achievement in grades 4 and 8 math and reading, even when eligibility for free or reduced price lunch and urban location were factored into the comparison. When minority status was used as a factor, it was found that black and Hispanic charter school students scored lower in 4th grade math and reading, but the difference was not significant. The fact that virtual charter school students were not shown to score lower than nonvirtual school students in this meta-analysis is an indicator of the success of distance education for K–12 learning.

Teacher quality and classification as a charter school have been recognized as factors that can influence student learning in classrooms, but little data is available about the influences of these factors in virtual schooling. Practitioners and policymakers in K–12 distance education are urged to use data-driven decision making, and to do so they must be informed by experience and data must be available. In 2004, there have been fewer than ten years of accumulated experience and too little detailed research published on web-based distance education methods. 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.

Conclusions

Students can experience similar levels of academic success while learning using telecommunications and learning in classroom settings. While distance learning as it is practiced in today's virtual schools uses technology that is less than ten years old and advances rapidly, the literature has shown that a student's education online can be as effective as it is in a classroom, provided that a classroom with the appropriate course is accessible to the student. As the power of communication technology and educational technology grow, the skill of distance educators and designers will be challenged to provide experiences that use that power to provide an experience for students that improves on classroom instruction with its limits of time and place. Research in K–12 distance education is maturing alongside the technology and those who use it, but current web-based distance education systems have only been studied for about the last five years at the K–12 level, a very short time in which to build a body of literature.

This meta-analysis represents an investigation into the literature on K–12 web-based distance education with attention on the factors likely to influence student performance. The result shows variation in the degree of success students have experienced, and a need for more information if firm conclusions are to be drawn. Blomeyer (2002) stated the recommendation well: "Support for additional professionally designed and executed program evaluations and scientific educational research should be given a high priority in all public and private agencies supporting effective implementation and use of online learning in K–12 learning communities" (page 10).

The importance of knowledge about effective virtual schooling cannot be overstated, because of the current boom in the numbers of virtual schools and students, and because of the essential role virtual schools can play in school reform movements and workforce development efforts. As of spring, 2004, there were roughly 2,400 publicly-funded cyber-based charter schools and state and district virtual schools in 37 U.S. states, with an estimated 40,000 to 50,000 students participating in online courses, according to Susan Patrick, Director of the U.S. Department of Education's Office of Educational Technology (Fording, 2004). With recent and continued growth in virtual schools, virtual school leaders and policy makers will need a strong research foundation on which to base decisions.

Several groups in the U.S. have identified school reform, particularly high school reform, as priorities in coming years. The U.S. Department of Education has identified high school reform models that support student achievement, and has recognized small school size, scheduling choice, charter schools, career academies, early college initiatives, and student engagement as research-based models that contribute to improved student achievement (U.S. Department of Education, 2004). The National Governor's Association has formed a task force to study redesigning high schools in order to make them "more rigorous and relevant to the lives of America's youth" (National Governor's Association, 2004). The task force initiative responds to employers' needs for more skilled and better educated workers by suggesting that reforms include choices in high school programs and opportunities to earn college credit or professional credentials. The National Association of Secondary School Principals in 2004 published *Breaking Ranks II*, which calls for reforming high schools to become more rigorous and personalized (National Association of Secondary School Principals, 2004), and the National High School Alliance has developed the *Catalog of Research on Secondary School Reform*. The catalog compiles studies of effective school reform programs, including those based on early college, smaller schools, student interests and learning styles, at-risk student needs, talent development, and career academies (National High School Alliance, 2004). Each of the reform models described and recommended by these groups is an example of a strength that has been shown by virtual schools. By offering scheduling flexibility, personalization, freedom from a large physical school, engaging tools of distance learning, opportunities to accelerate learning, and access to rigorous academic programs, virtual schools are not just important examples of school reform models, but virtual schools may represent the best hope for bringing high school reform quickly to large numbers of students.

Another strength of virtual schools is their unique capability for immersing students in information and communication technologies (ICT). An international effort is underway to improve ICT literacy as a "contribution to the development of human capital" (Educational Testing Service, 2001, p. iii). An international panel convened by the Educational Testing

Service determined that ICT skill is needed by citizens to function in the current technological climate, and that ICT skills are needed to help people worldwide meet fundamental needs, making ICT literacy a global objective. The development of ICT literacy begins with access to technology, and many publicly-funded virtual schools have found ways to bridge the access divide by providing computers to students. Virtual school students must develop ICT skills to be successful in online learning, and they may become the sought-after employees of the near future. Because of the global need for ICT skills and their role in virtual schools, demand could rise for data on effective virtual schools as more are developed worldwide.

The Need for Prospective Study in Virtual Schooling

An important step toward improving the state on virtual schools research was taken in 2004 when the U.S. Department of Education hosted an E-learning Summit to explore the status of K–12 e-learning in the U.S. The DOE Office of Educational Technology is showing leadership by identifying e-learning as a priority in the new National Educational Technology Plan. Technology, including e-learning, is seen as a force that can transform education because of the power of e-learning to individualize, personalize and differentiate instruction. Plans for the Federal role in e-learning leadership will include development of an e-learning clearinghouse listing programs for students, a process for addressing quality and accreditation issues, and support for developing online content. Such initiatives begin to bring knowledge and expertise to more stakeholders, assist policymakers and practitioners in accessing information, and serve as a focal point for guiding future work that will improve outcomes across the spectrum.

As a relatively recent innovation in the sometimes slow-moving world of education, distance education has been shown over decades with every variety of technology to work effectively although it works in very different ways than classroom instruction does, it meets different needs, and serves different audiences, having had far less time in which to mature, as evidenced by the studies included in this meta-analysis. The literature contains reports on distance education programs in which student outcomes exceed those in conventional classrooms (see citations in “Distance Education in the K–12 Context” section), but in order to make use of such data in syntheses such as this one, complete data need to be reported.

Recommendations for K–12 Online Learning Policy and Practice

Policy-makers and practitioners should continue to move forward in developing and implementing K–12 distance education programs when those programs meet identified needs and when they are designed and managed as carefully as traditional education programs. The “no significant difference” result reported here and elsewhere lends confidence to distance educators that their ongoing efforts are likely to be as effective as classroom-based education. This synthesis, considered together with current policy and recent research findings, demonstrates that students of many types and ages can learn in many content areas using the flexibility and choices afforded by distance education. In their recent article, *New Millenium Research for Educational Technology: A Call for a National Research Agenda,* Roblyer and Knezek (2003) recommended a focus and priorities for a future technology research agenda. The focus, they stated, should be providing a rationale for technology use. The priority is to explain why students and educators should use technology.

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 small subset of educational researchers have elected to focus on virtual schooling, either as doctoral candidates, faculty, program directors or independent evaluators.
- 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. Distance education has been seen primarily as a substitute for classroom instruction, rather than a potentially more effective way of learning. Until the goal is established of reaching a higher potential, research will continue to determine whether distance education is as effective as classroom instruction, rather than looking for ways that distance education can excel.
- 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. Most reports on virtual schooling released in the past omit sample sizes, mean scores, standard deviations, and other details needed for big-picture 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.

   This crucial informational campaign requires professionals working in distance education in any capacity to network by participating in conferences, publishing articles and papers, and contributing to discussions locally and globally where people who are not involved in distance education can learn.

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.

   The list of potential benefits should be broad, and should be a close match to the benefits or “effects” anticipated for any educational experience. Curriculum content should include a liberal education in which knowledge, skills, and dispositions are developed that successful students need in order to enjoy a full life in a democracy. But effects and benefits should also include academic literacies, technology skills, and academic standards.

This list of crucial, performance-based knowledge, skills, and dispositions must serve as a guide in the stages of design, implementation, and evaluation of programs. Consensus is needed on the goals of distance education, and plans should follow to evaluate progress toward those goals. Distance education program directors should see researchers as partners in informing practice.

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. A descriptive system supporting comparative analysis of all varieties of traditional and online and distance learning delivery systems will dramatically increase both the generalizability of results and the synthesizability of research findings available to inform development, implementation and institutionalization of online and distance learning programs.

4. Finally, 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 professional organizations like the North American Council on Online Learning and International Society for Technology in Education. The United States Department of Education and the leading professional organizations and groups should assume a leadership role organizing a national distance learning and online learning community of practice to work toward enacting these *essential* action recommendations.

   Distance educators belong to a wide variety of overlapping professional groups and associations that have the potential to contribute to a powerful and effective coalition. The larger coalition needed to weld a broader professional consensus should serve as a central clearinghouse for information about K–12 online and distance education, a matchmaking service for programs and evaluators, and as an organizational focus for organizing national efforts to support online and distance learning policy, program development, and professional development.

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.

References

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

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *The American Journal of Distance Education, 16,* 83–97.

*Alaska Department of Education and Early Development. (2003). *Spring 2003 benchmarks, Delta/Greely school district - Delta Charter Cyber School.* Retrieved July 19, 2004, from http://www.eed.state.ak.us/tls/assessment/AsmtVer/SchoolAsmtVerSuptSearch.cfm

Arkansas Virtual School. (2003). *2002–2003 testing overview*. Dardanelle, AR: author.

Bangert-Drowns, R., Hurley, M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74* (1), 29–58.

Barker, K., & Wendel, T. (2001). *E-learning: studying Canada's virtual secondary schools*. Kelowna, BC: Society for the Advancement of Excellence in Education. Retrieved April 11, 2004, from http://www.excellenceineducation.ca/pdfs/006.pdf

Bernard, R. M., Lou, Y., Abrami, P.C., Wozney, L., Borokhovski, E., Wallet, P.A., Wade, A., & Fiset, M. (2003, April). *How does distance education compare to classroom instruction? A meta-analysis of the empirical literature*. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago, IL.

Bianchi, W. (2002). The Wisconsin school of the air: Success story with implications. *Educational Technology & Society 5* (1).

Blomeyer, R. (2002). Virtual schools and e-learning in K–12 environments: Emerging policy and practice. *NCREL Policy Issues,* 11. Naperville, IL: North Central Regional Educational Laboratory. Retrieved May 2, 2004, from http://www.ncrel.org/policy/pubs/pdfs/pivol11.pdf

Bolton, J. (2002). *Web-based distance education: Pedagogy, epistemology, and instructional design.* University of Saskatchewan. Retrieved April 21, 2004, from http://www.usask.ca/education/coursework/802papers/boulton/boulton.pdf

Bond, A. (2002). *Learning music online: An accessible program for isolated students*. Kensington Park, SA: Australian National Training Authority. Retrieved April 12, 2004, from 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.

Brooks, J. & Brooks, M. (1993). *The case for constructivist classrooms*. Alexandria, VA: Association for Supervision and Curriculum Development.

Bruckman, A. (1998). Computer supported cooperative work. *The Journal of Collaborative Computing, 7,* 47–86.

Calderoni, J. (1998). *Telesecundaria: Using TV to bring education to rural Mexico*. Education and Technology Technical Notes Series: World Bank Human Development Network. Retrieved April 8, 2004, from http://wbln0018.worldbank.org/HDNet/HDdocs.nsf/C11FBFF6C1B77F9985256686006DC949/1635F1703FE053B385256754006D8C3F/$FILE/telesecundaria.pdf

Cavanaugh, C. S. (2001). The effectiveness of interactive distance education technologies in K–12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7(*1), 73–88.

* Colorado Department of Education. (2003a). *Branson alternative school, grades 1–6 school accountability report.* Denver, CO: Author. Retrieved July 19, 2004, from 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. Retrieved July 19, 2004, from 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. Retrieved July 19, 2004, from 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. Retrieved July 19, 2004, from http://reportcard.cde.state.co.us/reportcard/pdf/2003_0880_1887_M.pdf

Conzemius, A; & Sandrock, P. (2003). *Developing world language programs in elementary grades.* Madison, WI: Wisconsin State Department of Public Instruction. (ERIC Document Reproduction Service No. ED 480156)

Darling-Hammond, L. (2000). Teacher quality and student achievement: a review of state policy evidence. *Education Policy Analysis Archives, (8)*1.

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.

Duffy, T., & Cunningham, D. (1996). Constructivism: Implications for the design and delivery of instruction. In Jonassen, D. (Ed.), *Handbook of research on educational telecommunications and technology*. New York: Macmillan Library Reference.

Duffy, T., & Jonassen, D. (Eds.). (1992). *Constructivism and the technology of instruction: A conversation.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publisher.

Educational Testing Service. (2001). *Digital transformation: A framework for ICT literacy*. Retrieved October 9, 2004, from http://www.ets.org/research/ictliteracy/ictreport.pdf

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

Fording, L. (2004). Education, 21st century-style. *Newsweek*. Retrieved September 9, 2004, from http://www.msnbc.msn.com/id/4633126/

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., & Woods, M. (2003). Wellness NutriFit online learning in physical education for high school students. *Journal of Interactive Online Learning, 2* (2). Retrieved April 9, 2004, from http://www.ncolr.org/jiol/archives/2003/fall/03/index.html

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

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.

* Indiana Department of Education. (2004). *School snapshot: Irvington community school.* Indianapolis, IN: Author. Retrieved July 20, 2004, from 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., Zucker, A., Espinoza, C., McGhee, R., Yarnall, L., Zalles, D., & Lewis, A. (2000). *The online course experience: evaluation of the virtual high school's third year of implementation, 1999-2000*. Arlington, VA: SRI. Retrieved December 8, 2003, from http://www.govhs.org/Images/SRIEvals/$file/SRIAnnualReport2000.pdf

Kromrey, J., & Hogarty, K. (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.

Lipsey, B., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE Publications, Inc.

Machtmes, K., & Asher, J. W. (2000). A meta-analysis of the effectiveness of telecourses in distance education. *The American Journal of Distance Education, 14* (1), 27–46.

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. Retrieved April 12, 2004, from http://www.helsinki.fi/science/optek/1994/n4/mcgreal.txt

Means, B., & Haertel, G. (2004). A blueprint for a national research agenda to evaluate educational technology. In Means, B. & Haertel, G. (Eds.). *Using Technology Evaluation to Enhance Student Learning.* 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. Retrieved May 2, 2004, from http://media.lsi.ku.edu/research/vgeval/vgevaluationreport.pdf

* Minnesota Department of Education. (2003). *2003 report card: cyber village academy*. Minneapolis, MN: Author. Retrieved July 21 2004, from 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. Retrieved April 29, 2004, from http://hobbes.lite.msu.edu/~robmock/masters/abstract.htm

National Association of Secondary School Principals. (2004). *Breaking ranks II: Strategies for leading high school reform*. Retrieved October 9, 2004, from http://www.nassp.org/breakingranks/breakingransk2.cfm

National Association of State Boards of Education. (2001). *Any time, any place, any path, any pace: Taking the lead on e-learning policy.* Alexandria, VA: author.

National Governor's Association. (2004). *Governors leading high school reform efforts.* Retrieved October 9, 2004, from http://www.nga.org/nga/legislativeUpdate/1,1169,C_ISSUE_BRIEF^D_7342,00.html

National High School Alliance. (2004). *Catalog of research on secondary school reform.* Retrieved October 9, 2004, from http://www.hsalliance.org/research/index.asp

Nelson, F., Rosenburg, B. & Van Meter, N. (2004). *Charter school achievement on the 2003 national assessment of educational progress.* Washington, DC: American Federation of Teachers. Retrieved September 9, 2004, from http://www.aft.org/pubs-reports/downloads/teachers/NAEPCharterSchoolReport.pdf

Roblyer, M. & Knezek, G. (2003). New millennium research for educational technology: A call for a national research agenda. *Journal of Research on Technology in Education, 36* (1), 60–71.

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. Retrieved April 13, 2004, from http://www.cade-aced.ca/en_pub.php

Sabelli, N. (2004). Policy, planning, and the evaluation of learning technology. In Means, B. & Haertel, G. (Eds.). *Using Technology Evaluation to Enhance Student Learning*. New York: Teachers College Press.

* Schollie, B. (2001). *Student achievement and performance levels in online education research study.* Edmonton, Alberta: Alberta Online Consortium. Retrieved April 25, 2004, from http://www.albertaonline.ab.ca/pdfs/AOCresearch_full_report.pdf

Shachar, M. & Neumann, Y. (2003). Differences between traditional and distance education academic performances: a meta-analytic approach. *International Review of Research in Open and Distance Learning*. October.

Slavin, R. (2003). *Educational psychology theory and practice*, 7th Ed. Boston, MA: Pearson Education, Inc.

Southern Regional Education Board. (2003). *Essential principles of high-quality online teaching: guidelines for evaluating K–12 online teachers*. Atlanta, GA: author. Retrieved May 2, 2004, from http://www.sreb.org/programs/EdTech/pubs/PDF/Essential_Principles.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*. Retrieved April 9, 2004, from http://www.tellearn.mun.ca/pubs/virtual.html

* Texas Education Agency. (2003). *2002–2003 campus performance: Southwest Virtual Preparatory School.* Austin, TX: Author. Retrieved July 22, 2004, from http://www.tea.state.tx.us/perfreport/aeis/2003/campus.srch.html

Ungerleider, C.S., & Burns, T.C. (2003). *A systematic review of the effectiveness and efficiency of networked ICT in education: A state of the field report to the Council of Ministers of Education, Canada and Industry Canada*. Unpublished report.

U.S. Department of Education. (2004). *High school reform models that support student achievement.* Retrieved October 9, 2004, from http://www.ed.gov/about/offices/list/ovae/pi/hs/reform.html

* Washington Office of the Superintendent of Public Instruction. (2003). *Internet academy 2002–2003 WASL results.* Olympia, WA: Author. Retrieved July 22, 2004, from http://reportcard.ospi.k12.wa.us/Reports/summary.aspx?schoolId=1164&reportLevel=School

Wedemeyer, C.A. (1981). *Learning at the back door: Reflections on non traditional learning in the lifespan.* Madison: University of Wisconsin Press.

* Wisconsin Department of Public Instruction. (2003). *Successful school guide: Wisconsin Virtual Academy.* Madison, WI: Author. Retrieved July 22, 2004, from http://data.dpi.state.wi.us/data/questions.asp?fullkey=011945040100&DN=Northern+Ozaukee&SN=Wisconsin+Virtual+Academy&TYPECODE=6&CTY=45&ORGLEVEL=SC

Yasin, K. & Luberisse, Y. (1997). *Meeting the needs of a new democracy: Multichannel learning and interactive radio instruction in Haiti: a case study*. Washington, DC: USAID. Retrieved April 29, 2004, from http://ies.edc.org/pubs/book11.htm

Appendix

Coded Variables and Study Features in the Codebook

A. Identification of studies

1. Study number ("study").

2. Finding/hypothesis number ("finding").

3. Author name ("author"). Last name of first author.

4. Year of publication ("year").

5. Number of findings/hypotheses within study ("number").

6. Country ("country").
   Unspecified=0,
   USA=1,
   Canada=2,
   Mexico/Central America/South America=3,
   Europe=4,
   Asia=5,
   Africa=6,
   Australia/Pacific=7,
   Multinational=8,
   Other=9.

7. Grade level of students ("grade").
   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.

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

9. Content area ("content").
Unspecified=0,
Reading/language arts=1,
Mathematics=2,
Social studies=3,
Science=4,
Computers/technology=5,
Foreign language=6,
Arts=7,
Physical education=8,
Other=9,
Writing=10.

10. Type of publication ("publication").
Published journal article=1,
Journal article in press=2,
Book chapter=3,
Report=4,
Dissertation=5,
Conference paper=6.

B. Distance learning features

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

2. Frequency of distance learning experience ("frequency").
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.

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

4. Number of distance learning sessions ("dlnumber").
Unspecified=0,
List number of sessions.

5. Duration of distance learning sessions ("dlduration").
Unspecified=0,
List average minutes per session.

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

7. Instructor role ("instructrole").
Unspecified=0,
Fully moderated=1,
Nonmoderated=2,
Combination=3,
Other=4.

8. Timing of interactions ("timing").
Unspecified=0,
Synchronous=1,
Asynchronous=2,
Combination=3,
Other =4.

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

C. Instructor/program features

1. Amount of teacher preparation in distance learning ("instructprep").
Unspecified=0,
List hours of preparation.

2. Amount of teacher experience in distance learning ("instructexp").
Unspecified=0,
List years of experience.

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

4. Setting of students during distance learning (“setting”).
   Unspecified=0,
   Home=1,
   School=2,
   Other=3,
   Combination=4.

D. Study quality features

1. Student sample size (“sample”). Actual sample size.

2. Measure of academic outcome (“achmeasure”).
   Standardized test=1,
   Researcher-made test=2,
   Teacher-made test=3,
   Other=4.

3. Testing sequence (“testseq”).
   Unspecified=0,
   Pre-post=1,
   Post only=2,
   Other=3.

4. Pretest equivalency (“preequiv”). Have the initial differences between groups been accounted for?
   Unspecified=0,
   Statistical control (ANCOVA, regression)=1,
   Random assignment=2,
   Statistical control and random assignment=3,
   Gain scores=4,
   Other=5.

5. Reported reliability of measures (“reliability”).
   Unspecified=00,
   Actual reliability statistic.

6. Effect size coefficient (“effsize”).
   Actual coefficient.

7. Statistics used in determining effect size. ("esstats").
    Means=1,
    t-value=2,
    F-value=3,
    Chi-square=4,
    Other=5.

8. Weight ("weight").
    One divided by the actual number of findings/hypotheses in the study.

E. Sources of Invalidity

1. Type of Design ("design").
   Quasi-experimental/nonrandomized one group pretest-posttest=1,
   Nonrandomized static-group comparison=2,
   Nonrandomized pre-post control group=3,
   Time series=4,
   Randomized posttest-only control group=5,
   Randomized pre-post control group=6,
   Other=7.

2. History ("history"). Control for specific events occurring between the first and second measurement in addition to the experimental variable.
    Adequately controlled by design=1,
    Definite weakness of design=2,
    Possible source of concern=3,
    Not a relevant factor=4.

3. Maturation ("maturation"). Control for processes within the participants operating as a function of the passage of time.
    Are there processes within participants operating as a function of the passage of time, such as growing older or more tired, that might account for changes in the dependent measure?
    Adequately controlled by design=1,
    Definite weakness of design=2,
    Possible source of concern=3,
    Not a relevant factor=4.

4. Testing ("testing"). Control for the effect of taking a test upon the scores of a second testing.
    Adequately controlled by design=1,
    Definite weakness of design=2,
    Possible source of concern=3,
    Not a relevant factor=4.

5. Instrumentation (“instrument”). Control for changes in calibration or observers' scores that produce changes in the obtained measurement.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

6. Statistical Regression (“regression”). Control for group selection based on their extreme scores.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

7. Selection Bias (“selection”). Control for biases resulting in the differential selection of comparison groups.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

8. Mortality (“mortality”). Control for differential loss of participants from the experimental and control groups.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

9. Selection-Maturation Interaction (“selectmatur”). Control for interaction between extraneous factors such as history, maturation, or testing and the specific selection differences that distinguish the experimental and control groups.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

10. Reactive or Interaction Effect of Testing (“testeff”). Control for the influence of pre-testing on the participants' responsiveness to the experimental variable, making the results for a pre-tested population unrepresentative of the effects of the experimental variable for the unpre-tested universe from which the participants were selected.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

11. Interaction of Selection Biases and Treatment (“biastreat”). Control for selective factors upon which sampling was based which interact differentially with the experimental variable.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

12. Reactive Effects of Experimental Arrangements (“effexper”). Control for effects of the experimental setting that would preclude generalizing about the effect of the experimental variable upon persons being exposed to it in nonexperimental settings.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

13. Multiple-Treatment Interference (“multtreat”) Control for nonerasable effects of previous treatments applied to the same participants.
   Adequately controlled by design=1,
   Definite weakness of design=2,
   Possible source of concern=3,
   Not a relevant factor=4.

14. Statistical Power (“statpower”). Large enough sample size to reject the null hypothesis at a given level of probability, or estimate coefficients within reasonably small margins of error. A sample of over 60 for groups such as classes or schools; a sample of over 100 for individuals.
   Probable threat (<60 for groups or <100 for individuals as the unit of analysis)=1,
   Adequately minimized (>60 for groups, >100 for individuals)=2.