Evaluation of Evidence-Based Practices in Online Learning
A Meta-Analysis and Review of Online Learning Studies
Executive Summary

Online learning—both for students and for teachers—is one of the fastest growing trends in educational uses of technology. The National Center for Education Statistics (2008) estimated that the number of K–12 public school students enrolling in a technology-based distance education course grew by 65 percent in the two years from 2002-03 to 2004-05. On the basis of a more recent district survey, Picciano and Seaman (2009) estimated that more than a million K–12 students took online courses in school year 2007–08.

Online learning overlaps with the broader category of distance learning, which encompasses earlier technologies such as correspondence courses, educational television and videoconferencing. Earlier studies of distance learning concluded that these technologies were not significantly different from regular classroom learning in terms of effectiveness. Policy-makers reasoned that if online instruction is no worse than traditional instruction in terms of student outcomes, then online education initiatives could be justified on the basis of cost efficiency or need to provide access to learners in settings where face-to-face instruction is not feasible. The question of the relative efficacy of online and face-to-face instruction needs to be revisited, however, in light of today’s online learning applications, which can take advantage of a wide range of Web resources, including not only multimedia but also Web-based applications and new collaboration technologies. These forms of online learning are a far cry from the televised broadcasts and videoconferencing that characterized earlier generations of distance education. Moreover, interest in hybrid approaches that blend in-class and online activities is increasing. Policy-makers and practitioners want to know about the effectiveness of Internet-based, interactive online learning approaches and need information about the conditions under which online learning is effective.

The findings presented here are derived from (a) a systematic search for empirical studies of the effectiveness of online learning and (b) a meta-analysis of those studies from which effect sizes that contrasted online and face-to-face instruction could be extracted or estimated. A narrative summary of studies comparing different forms of online learning is also provided.

These activities were undertaken to address four research questions:

1. How does the effectiveness of online learning compare with that of face-to-face instruction?
2. Does supplementing face-to-face instruction with online instruction enhance learning?
3. What practices are associated with more effective online learning?
4. What conditions influence the effectiveness of online learning?

This meta-analysis and review of empirical online learning research are part of a broader study of practices in online learning being conducted by SRI International for the Policy and Program Studies Service of the U.S. Department of Education. The goal of the study as a whole is to provide policy-makers, administrators and educators with research-based guidance about how to implement online learning for K–12 education and teacher preparation. An unexpected finding of the literature search, however, was the small number of published studies contrasting online and
face-to-face learning conditions for K–12 students. Because the search encompassed the research literature not only on K–12 education but also on career technology, medical and higher education, as well as corporate and military training, it yielded enough studies with older learners to justify a quantitative meta-analysis. Thus, analytic findings with implications for K–12 learning are reported here, but caution is required in generalizing to the K–12 population because the results are derived for the most part from studies in other settings (e.g., medical training, higher education).

This literature review and meta-analysis differ from recent meta-analyses of distance learning in that they

- Limit the search to studies of Web-based instruction (i.e., eliminating studies of video- and audio-based telecourses or stand-alone, computer-based instruction);
- Include only studies with random-assignment or controlled quasi-experimental designs; and
- Examine effects only for objective measures of student learning (e.g., discarding effects for student or teacher perceptions of learning or course quality, student affect, etc.).

This analysis and review distinguish between instruction that is offered entirely online and instruction that combines online and face-to-face elements. The first of the alternatives to classroom-based instruction, entirely online instruction, is attractive on the basis of cost and convenience as long as it is as effective as classroom instruction. The second alternative, which the online learning field generally refers to as blended or hybrid learning, needs to be more effective than conventional face-to-face instruction to justify the additional time and costs it entails. Because the evaluation criteria for the two types of learning differ, this meta-analysis presents separate estimates of mean effect size for the two subsets of studies.

**Literature Search**

The most unexpected finding was that an extensive initial search of the published literature from 1996 through 2006 found no experimental or controlled quasi-experimental studies that both compared the learning effectiveness of online and face-to-face instruction for K–12 students and provided sufficient data for inclusion in a meta-analysis. A subsequent search extended the time frame for studies through July 2008.

The computerized searches of online databases and citations in prior meta-analyses of distance learning as well as a manual search of the last three years of key journals returned 1,132 abstracts. In two stages of screening of the abstracts and full texts of the articles, 176 online learning research studies published between 1996 and 2008 were identified that used an experimental or quasi-experimental design and objectively measured student learning outcomes. Of these 176 studies, 99 had at least one contrast between an included online or blended learning condition and face-to-face (offline) instruction that potentially could be used in the quantitative meta-analysis. Just nine of these 99 involved K–12 learners. The 77 studies without a face-to-face condition compared different variations of online learning (without a face-to-face control condition) and were set aside for narrative synthesis.
Meta-Analysis

Meta-analysis is a technique for combining the results of multiple experiments or quasi-experiments to obtain a composite estimate of the size of the effect. The result of each experiment is expressed as an effect size, which is the difference between the mean for the treatment group and the mean for the control group, divided by the pooled standard deviation. Of the 99 studies comparing online and face-to-face conditions, 45 provided sufficient data to compute or estimate 50 independent effect sizes (some studies included more than one effect). Four of the nine studies involving K–12 learners were excluded from the meta-analysis: Two were quasi-experiments without statistical control for preexisting group differences; the other two failed to provide sufficient information to support computation of an effect size.

Most of the articles containing the 50 effects in the meta-analysis were published in 2004 or more recently. The split between studies of purely online learning and those contrasting blended online/faceto-face conditions against face-to-face instruction was fairly even, with 27 effects in the first category and 23 in the second. The 50 estimated effect sizes included seven contrasts from five studies conducted with K–12 learners—two from eighth-grade students in social studies classes, one for eighth- and ninth-grade students taking Algebra I, two from a study of middle school students taking Spanish, one for fifth-grade students in science classes in Taiwan, and one from elementary-age students in special education classes. The types of learners in the remaining studies were about evenly split between college or community college students and graduate students or adults receiving professional training. All but two of the studies involved formal instruction. The most common subject matter was medicine or health care. Other content types were computer science, teacher education, mathematics, languages, science, social science, and business. Among the 48 contrasts from studies that indicated the time period over which instruction occurred, 19 involved instructional time frames of less than a month, and the remainder involved longer periods. In terms of instructional features, the online learning conditions in these studies were less likely to be instructor-directed (8 contrasts) than they were to be student-directed, independent learning (17 contrasts) or interactive and collaborative in nature (22 contrasts).

Effect sizes were computed or estimated for this final set of 50 contrasts. Among the 50 individual study effects, 11 were significantly positive, favoring the online or hybrid learning condition. Three contrasts found a statistically significant effect favoring the traditional face-to-face condition.1

---

1 When a \( \alpha < .05 \) level of significance is used for contrasts, one would expect approximately 1 in 20 contrasts to show a significant difference by chance. For 50 contrasts, then, one would expect 2 or 3 significant differences by chance. The finding of 3 significant contrasts associated with face-to-face instruction is within the range one would expect by chance; the 11 contrasts associated with online or hybrid instruction exceeds what one would expect by chance.
Narrative Synthesis

In addition to the meta-analysis comparing online learning conditions with face-to-face instruction, analysts reviewed and summarized experimental and quasi-experimental studies contrasting different versions of online learning. Some of these studies contrasted purely online learning conditions with classes that combined online and face-to-face interactions. Others explored online learning with and without elements such as video, online quizzes, assigned groups, or guidance for online activities. Five of these studies involved K–12 learners.

Key Findings

The main finding from the literature review was that

- Few rigorous research studies of the effectiveness of online learning for K–12 students have been published. A systematic search of the research literature from 1994 through 2006 found no experimental or controlled quasi-experimental studies comparing the learning effects of online versus face-to-face instruction for K–12 students that provide sufficient data to compute an effect size. A subsequent search that expanded the timeframe through July 2008 identified just five published studies meeting meta-analysis criteria.

The meta-analysis of 50 study effects, 43 of which were drawn from research with older learners, found that

- Students in online conditions performed modestly better, on average, than those learning the same material through traditional face-to-face instruction. Learning outcomes for students who engaged in online learning exceeded those of students receiving face-to-face instruction, with an average effect size of +0.20 favoring online conditions. The mean difference between online and face-to-face conditions across the 50 contrasts is statistically significant at the $p < .001$ level. Interpretations of this result, however, should take into consideration the fact that online and face-to-face conditions generally differed on multiple dimensions, including the amount of time that learners spent on task. The advantages observed for online learning conditions therefore may be the product of aspects of those treatment conditions other than the instructional delivery medium per se.

---

2 The meta-analysis was run also with just the 43 studies with older learners. Results were very similar to those for the meta-analysis including all 50 contrasts. Variations in findings when K-12 studies are removed are described in footnotes.

3 The + sign indicates that the outcome for the treatment condition was larger than that for the control condition. A − sign before an effect estimate would indicate that students in the control condition had stronger outcomes than those in the treatment condition. Cohen (1992) suggests that effect sizes of .20 can be considered “small,” those of approximately .50 “medium,” and those of .80 or greater “large.”

4 The $p$-value represents the likelihood that an effect of this size or larger will be found by chance if the two populations under comparison do not differ. A $p$-value of less than .05 indicates that there is less than 1 chance in 20 that a difference of the observed size would be found for samples drawn from populations that do not differ.
Instruction combining online and face-to-face elements had a larger advantage relative to purely face-to-face instruction than did purely online instruction. The mean effect size in studies comparing blended with face-to-face instruction was +0.35, \( p < .001 \). This effect size is larger than that for studies comparing purely online and purely face-to-face conditions, which had an average effect size of +0.05, \( p = .46 \). In fact, the learning outcomes for students in purely online conditions and those for students in purely face-to-face conditions were statistically equivalent. An important issue to keep in mind in reviewing these findings is that many studies did not attempt to equate (a) all the curriculum materials, (b) aspects of pedagogy and (c) learning time in the treatment and control conditions. Indeed, some authors asserted that it would be impossible to have done so. Hence, the observed advantage for blended learning conditions is not necessarily rooted in the media used per se and may reflect differences in content, pedagogy and learning time.

Effect sizes were larger for studies in which the online instruction was collaborative or instructor-directed than in those studies where online learners worked independently.\(^5\) The type of learning experience moderated the size of the online learning effect (\( Q = 6.19, p < .05 \)).\(^6\) The mean effect sizes for collaborative instruction (+0.25) and for instructor-directed instruction (+0.39) were significantly positive whereas the mean effect size for independent learning (+0.05) was not.

Most of the variations in the way in which different studies implemented online learning did not affect student learning outcomes significantly. Analysts examined 13 online learning practices as potential sources of variation in the effectiveness of online learning compared with face-to-face instruction. Of those variables, the two mentioned above (i.e., the use of a blended rather than a purely online approach and instructor-directed or collaborative rather than independent, self-directed instruction) were the only statistically significant influences on effectiveness. The other 11 online learning practice variables that were analyzed did not affect student learning significantly. However, the relatively small number of studies contrasting learning outcomes for online and face-to-face instruction that included information about any specific aspect of implementation impeded efforts to identify online instructional practices that affect learning outcomes.

The effectiveness of online learning approaches appears quite broad across different content and learner types. Online learning appeared to be an effective option for both undergraduates (mean effect of +0.30, \( p < .001 \)) and for graduate students and professionals (+0.10, \( p < .05 \)) in a wide range of academic and professional studies. Though positive, the mean effect size is not significant for the seven contrasts involving K–12 students, but the number of K–12 studies is too small to warrant much confidence in the mean effect estimate for this learner group. Three of the K–12 studies had

---

\(^5\) Online experiences in which students explored digital artifacts and controlled the specific material they wanted to view were categorized as independent learning experiences.

\(^6\) Online experiences in which students explored digital artifacts and controlled the specific material they wanted to view were categorized as “active” learning experiences. This contrast is not statistically significant (\( p = .13 \)) when the five K-12 studies are removed from the analysis.
significant effects favoring a blended learning condition, one had a significant negative
effect favoring face-to-face instruction, and three contrasts did not attain statistical
significance. The test for learner type as a moderator variable was nonsignificant. No
significant differences in effectiveness were found that related to the subject of
instruction.

- **Effect sizes were larger for studies in which the online and face-to-face conditions varied
  in terms of curriculum materials and aspects of instructional approach in addition to the
  medium of instruction.** Analysts examined the characteristics of the studies in the meta-
  analysis to ascertain whether features of the studies’ methodologies could account for
  obtained effects. Six methodological variables were tested as potential moderators: (a)
sample size, (b) type of knowledge tested, (c) strength of study design, (d) unit of
  assignment to condition, (e) instructor equivalence across conditions, and (f) equivalence
  of curriculum and instructional approach across conditions. Only equivalence of
  curriculum and instruction emerged as a significant moderator variable \( Q = 6.85, p < .01 \). Studies in which analysts judged the curriculum and instruction to be identical or
  almost identical in online and face-to-face conditions had smaller effects than those
  studies where the two conditions varied in terms of multiple aspects of instruction (+0.13
  compared with +0.40, respectively). Instruction could differ in terms of the way activities
  were organized (for example as group work in one condition and independent work in
  another) or in the inclusion of instructional resources (such as a simulation or instructor
  lectures) in one condition but not the other.

The narrative review of experimental and quasi-experimental studies contrasting different online
learning practices found that the majority of available studies suggest the following:

- **Blended and purely online learning conditions implemented within a single study
generally result in similar student learning outcomes.** When a study contrasts blended
  and purely online conditions, student learning is usually comparable across the two
  conditions.

- **Elements such as video or online quizzes do not appear to influence the amount that
students learn in online classes.** The research does not support the use of some frequently
  recommended online learning practices. Inclusion of more media in an online application
does not appear to enhance learning. The practice of providing online quizzes does not
  seem to be more effective than other tactics such as assigning homework.

- **Online learning can be enhanced by giving learners control of their interactions with
media and prompting learner reflection.** Studies indicate that manipulations that trigger
  learner activity or learner reflection and self-monitoring of understanding are effective
  when students pursue online learning as individuals.

- **Providing guidance for learning for groups of students appears less successful than does
using such mechanisms with individual learners.** When groups of students are learning
together online, support mechanisms such as guiding questions generally influence the
  way students interact, but not the amount they learn.
Conclusions

In recent experimental and quasi-experimental studies contrasting blends of online and face-to-face instruction with conventional face-to-face classes, blended instruction has been more effective, providing a rationale for the effort required to design and implement blended approaches. When used by itself, online learning appears to be as effective as conventional classroom instruction, but not more so.

However, several caveats are in order: Despite what appears to be strong support for blended learning applications, the studies in this meta-analysis do not demonstrate that online learning is superior as a medium. In many of the studies showing an advantage for blended learning, the online and classroom conditions differed in terms of time spent, curriculum and pedagogy. It was the combination of elements in the treatment conditions (which was likely to have included additional learning time and materials as well as additional opportunities for collaboration) that produced the observed learning advantages. At the same time, one should note that online learning is much more conducive to the expansion of learning time than is face-to-face instruction.

In addition, although the types of research designs used by the studies in the meta-analysis were strong (i.e., experimental or controlled quasi-experimental), many of the studies suffered from weaknesses such as small sample sizes; failure to report retention rates for students in the conditions being contrasted; and, in many cases, potential bias stemming from the authors’ dual roles as experimenters and instructors.

Finally, the great majority of estimated effect sizes in the meta-analysis are for undergraduate and older students, not elementary or secondary learners. Although this meta-analysis did not find a significant effect by learner type, when learners’ age groups are considered separately, the mean effect size is significantly positive for undergraduate and other older learners but not for K–12 students.

Another consideration is that various online learning implementation practices may have differing effectiveness for K–12 learners than they do for older students. It is certainly possible that younger online students could benefit from practices (such as embedding feedback, for example) that did not have a positive impact for college students and older learners. Without new random assignment or controlled quasi-experimental studies of the effects of online learning options for K–12 students, policy-makers will lack scientific evidence of the effectiveness of these emerging alternatives to face-to-face instruction.