The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature

BARBARA MEANS

SRI International

YUKIE TOYAMA

SRI International

ROBERT MURPHY

SRI International

MARIANNE BAKI

SRI International

Background/Context: *Earlier research on various forms of distance learning concluded that these technologies do not differ significantly from regular classroom instruction in terms of learning outcomes. Now that web-based learning has emerged as a major trend in both K–12 and higher education, the relative efficacy of online and face-to-face instruction needs to be revisited. The increased capabilities of web-based applications and collaboration technologies and the rise of blended learning models combining web-based and face-to-face classroom instruction have raised expectations for the effectiveness of online learning.*

Purpose/Objective/Research Question/Focus of Study: *This meta-analysis was designed to produce a statistical synthesis of studies contrasting learning outcomes for either fully online or blended learning conditions with those of face-to-face classroom instruction.*

Population/Participants/Subjects: *The types of learners in the meta-analysis studies were about evenly split between students in college or earlier years of education and learners in*

graduate programs or professional training. The average learner age in a study ranged from 13 to 44.

Intervention/Program/Practice: *The meta-analysis was conducted on 50 effects found in 45 studies contrasting a fully or partially online condition with a fully face-to-face instructional condition. Length of instruction varied across studies and exceeded one month in the majority of them.*

Research Design: *The meta-analysis corpus consisted of (1) experimental studies using random assignment and (2) quasi-experiments with statistical control for preexisting group differences. An effect size was calculated or estimated for each contrast, and average effect sizes were computed for fully online learning and for blended learning. A coding scheme was applied to classify each study in terms of a set of conditions, practices, and methodological variables.*

Findings/Results: *The meta-analysis found that, on average, students in online learning conditions performed modestly better than those receiving face-to-face instruction. The advantage over face-to-face classes was significant in those studies contrasting blended learning with traditional face-to-face instruction but not in those studies contrasting purely online with face-to-face conditions.*

Conclusions/Recommendations: *Studies using blended learning also tended to involve additional learning time, instructional resources, and course elements that encourage interactions among learners. This confounding leaves open the possibility that one or all of these other practice variables contributed to the particularly positive outcomes for blended learning. Further research and development on different blended learning models is warranted. Experimental research testing design principles for blending online and face-to-face instruction for different kinds of learners is needed.*

Online learning is one of the fastest growing trends in educational uses of technology. By the 2006–2007 academic year, 61% of US higher education institutions offered online courses (Parsad & Lewis, 2008). In fall 2008, over 4.6 million students—over one quarter of all U.S. higher education students—were taking at least one online course (Allen & Seaman, 2010). In the corporate world, according to a report by the American Society for Training and Development, about 33% of training was delivered electronically in 2007, nearly triple the rate in 2000 (Paradise, 2008).

Although K–12 school systems lagged behind other sectors in moving into online learning, this sector's adoption of e-learning is now proceeding rapidly. As of late 2009, 45 of the 50 states and Washington DC had at least one form of online program, such as a state virtual school offering courses to supplement conventional offerings in brick-and-mortar schools, a state-led online initiative, or a full-time online school (Watson, Gemin, Ryan, & Wicks, 2009). The largest state virtual school, the Florida Virtual School, had more than 150,000 course enrollments in 2008–2009. A number of states, including Michigan, Florida, Alabama, and Idaho,

have made successful completion of an online course a requirement for earning a high school diploma.

Two district surveys commissioned by the Sloan Consortium (Picciano & Seaman 2007; 2008) produced estimates that 700,000 K–12 public school students took online courses in 2005–2006, and more than a million students did so in 2007–2008: a 43% increase in just 2 years. Christensen, Horn, and Johnson (2008) predicted that by 2019, one half of all U.S. high school enrollments will be online.

Online learning has become popular because of its potential for providing more flexible access to content and instruction at any time, from any place. Frequently, the motivation for online learning programs entails (1) increasing the availability of learning experiences for learners who cannot or choose not to attend traditional face-to-face offerings, (2) assembling and disseminating instructional content more cost-efficiently, and/or (3) providing access to qualified instructors to learners in places where such instructors are not available. Online learning advocates argue further that additional reasons for embracing this medium of instruction include current technology's support of a degree of interactivity, social networking, collaboration, and reflection that can enhance learning relative to normal classroom conditions (Rudestam & Schoenholtz-Read, 2010).

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 reported overall effect sizes near zero, indicating that learning with these technologies, taken as a whole, was not significantly different from regular classroom learning in terms of effectiveness (Bernard et al., 2004; Cavanaugh, 2001; Machtmes & Asher, 2000; Zhao, Lei, Yan, & Tan, 2005). 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 the need to provide access to learners in settings where face-to-face instruction is not feasible (Florida Tax, 2007; Wise & Rothman, 2010). Research finding no difference in effectiveness does not justify moving instruction online in cases in which students have access to classroom instruction and cost savings are not expected. However, members of the distance education community view the advent of online, web-based learning as significantly different from prior forms of distance education, such as correspondence courses and one-way video. Online learning has been described as a "fifth generation" version of distance education "designed to capitalize on the features of the Internet and the Web" (Taylor, 2001, p. 2). Taylor concluded:

> Previous generations of distance education are essentially a function of resource allocation parameters based on the traditional cottage industry model, whereas the fifth generation based on automated response systems has the potential not only to improve economies of scale but also to improve the pedagogical quality and responsiveness of service to students. (p. 8)

The question of the relative efficacy of online and face-to-face instruction needs to be revisited in light of the advent of fifth-generation distance learning and today's online learning applications, which can take advantage of a wide range of web resources, including web-based applications (e.g., audio/video streaming, learning management systems, 3D simulations and visualizations, multiuser games) and new collaboration and communication technologies (e.g., Internet telephony, chat, wikis, blogs, screen sharing, shared graphical whiteboards). Learning that is supported by these Internet-based tools and resources is a far cry from the televised broadcasts and videoconferencing that characterized earlier generations of distance education. Online learning proponents suggest that these newer technologies support learning that is not just as good as, but better than, conventional classroom instruction (National Survey of Student Engagement, 2008; M.S. Smith, 2009; Zhao et al., 2005).

Learning technology researchers too see the Internet not just as a delivery medium but also as a potential means to enhance the quality of learning experiences and outcomes. One common conjecture is that learning a complex body of knowledge effectively requires a community of learners (Bransford, Brown, & Cocking, 1999; Riel & Polin, 2004; Schwen & Hara, 2004; Vrasidas & Glass, 2004) and that online technologies can be used to expand and support such communities, promoting "participatory" models of education (Barab, Squire, & Dueber, 2000; Barab & Thomas, 2001). Research in this area tends to be descriptive of individual learning systems, however, with relatively few rigorous empirical studies comparing learning outcomes for online and conventional approaches (Dynarski et al., 2007; R. Smith, Clark, & Blomeyer, 2005).

Another important trend in recent years is the emergence of "blended" or "hybrid" approaches that combine online activities and face-to-face instruction (Graham, 2005). As early as 2002, the president of Pennsylvania State University stated that "hybrid instruction is the single greatest unrecognized trend in higher education today" (Young, 2002, p. A33). Similarly, in 2003, the American Society for Training and Development identified blended learning as among the top 10 trends to emerge in the knowledge delivery industry (Rooney, 2003). In K–12 education, a recent study by the North American Council for Online

Learning predicted that the blended approach is likely to emerge as the predominant model of instruction and become far more common than either conventional, purely face-to-face classroom instruction or instruction done entirely online (Watson, 2008).

The terms *blended learning* and *hybrid learning* are used interchangeably and without a broadly accepted precise definition. Bonk and Graham (2005) described blended learning systems as a combination of face-to-face instruction and computer-mediated instruction. The 2003 Sloan Survey of Online Learning (Allen & Seaman, 2003) provided somewhat more detail, defining blended learning as a "course that is a blend of the online and face-to-face course. Substantial proportion of the content is delivered online, typically uses online discussions, typically has some face-to-face meetings" (p. 6). Horn and Staker (2010) defined blended learning as "any time a student learns at least in part in a supervised brick-and-mortar location away from home *and* at least in part through online delivery with some element of student control over time, place, path and/or pace" (p. 3).

Blended approaches do not eliminate the need for a face-to-face instructor and usually do not yield cost savings as purely online offerings do. To justify the additional time and costs required for developing and implementing blended learning, policy makers want evidence that blended learning is not just as effective as, but actually more effective than, traditional face-to-face instruction.
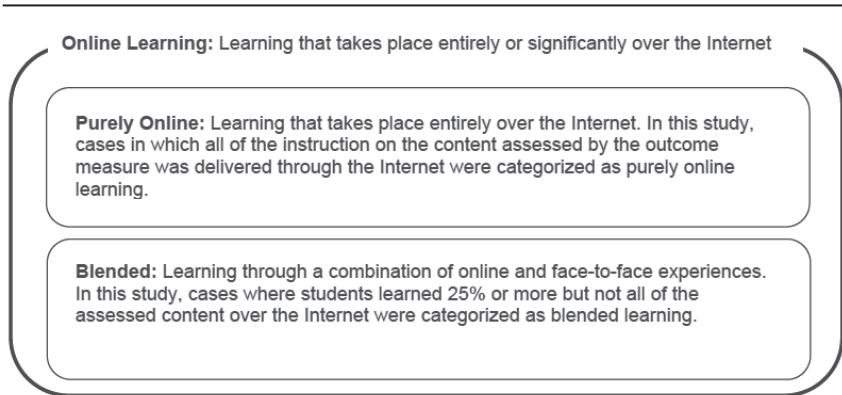
Further, for both blended and purely online learning, policy makers and practitioners need research-based information about the conditions under which online learning is effective and the practices associated with more effective online learning. The present article reports a meta-analytic study that investigated the effectiveness of online learning in general, and both purely online and blended versions of online learning in particular, for a variety of learners and with a range of different contexts and practices.

CONCEPTUAL FRAMEWORK

As noted, online and blended learning are not clearly defined in the literature. For the purpose of this meta-analysis, we adopted the Sloan Consortium's definition of online learning as learning that takes place entirely or in substantial portion over the Internet. We operationalized the concept of "substantial portion" as 25% or more of the instruction on the content assessed by a study's learning outcome measure. This criterion was used to avoid including studies of incidental uses of the Internet, such as downloading references and turning in assignments.[1] Cases in

which all or substantially all of the instruction on the content assessed in the outcome measure was conducted over the Internet were categorized as "purely online," whereas those in which 25% or more, but not all, of the instruction on the content to be assessed occurred online were classified as "blended." The relationships among online learning, purely online learning, and blended learning as defined in this study are illustrated in Figure 1.

Figure 1. Definitions of online learning

Online Learning: Learning that takes place entirely or significantly over the Internet

Purely Online: Learning that takes place entirely over the Internet. In this study, cases in which all of the instruction on the content assessed by the outcome measure was delivered through the Internet were categorized as purely online learning.

Blended: Learning through a combination of online and face-to-face experiences. In this study, cases where students learned 25% or more but not all of the assessed content over the Internet were categorized as blended learning.

Although our research questions focus on the effectiveness of purely online and blended learning, we recognize that different types of factors can affect the size and direction of differences in student learning outcomes when comparing online and face-to-face conditions. Online learning opportunities differ also in terms of the setting where they are accessed (classroom, home, informal), the nature of the content (both the subject area and the type of learning, such as fact, concept, procedure, or strategy), and the technology involved (e.g., audio/video streaming, Internet telephony, podcasting, chat, simulations, videoconferencing, shared graphical whiteboard, screen sharing). A review of the moderator variables included in prior meta-analyses (Bernard et al., 2004; Sitzmann, Kraiger, Stewart, & Wisher, 2006; Zhao et al., 2005) informed the development of a conceptual framework for the current meta-analysis. That framework includes not only online learning practices, as discussed, but also the conditions under which the study was conducted (e.g., the type of students and content involved) and features of the study method (e.g., experimental design, sample size).

This conceptual framework, shown in Figure 2, guided the coding of studies included in the meta-analysis. At the top level, we conceive the

types of variables that may influence effect sizes as those relating to the online instruction practices, conditions under which the study was conducted, and aspects of the study methodology. Within each of these categories, we specified a set of specific features that have been hypothesized or found to influence learning outcomes in prior meta-analyses of distance learning (Bernard et al., 2004; Sitzmann et al., 2006; Zhao et al., 2005).

Figure 2. Conceptual framework
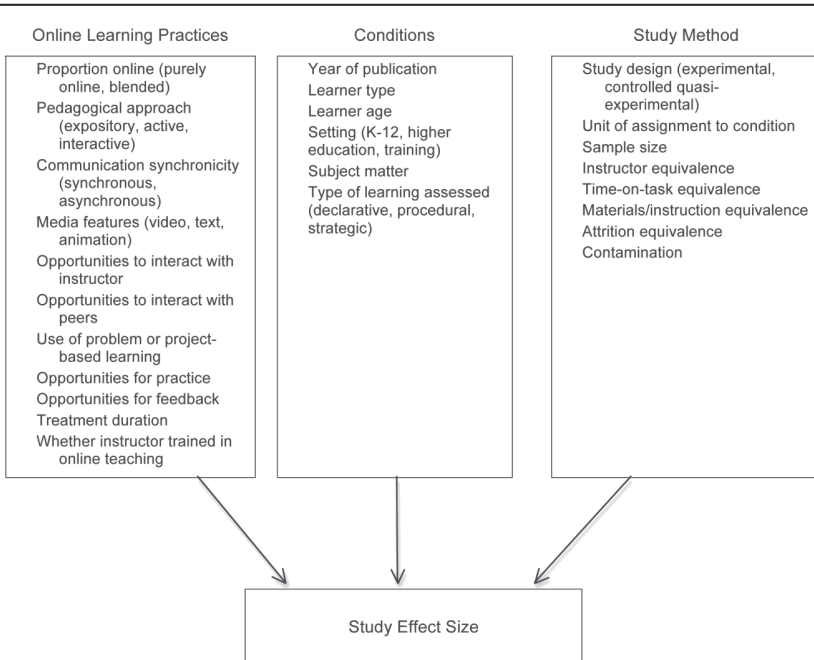
All these variables were coded, and in cases in which an adequate number of studies with the necessary information was available, a variable was tested as a potential moderator of the online learning effect size.

As discussed, from a practical perspective, one of the most fundamental distinctions among different online learning activities is whether they are blended or conducted purely online. Purely online instruction serves as a *replacement* for face-to-face instruction (e.g., a virtual course), with attendant implications for school staffing and cost savings. Purely online instruction may be an attractive alternative for cost reasons if it is equivalent to traditional face-to-face instruction in terms of student outcomes. Blended learning, on the other hand, is expected to be an *enhancement* of

face-to-face instruction. Many would consider blended learning applications that produce learning outcomes that are merely equivalent to (not better than) those resulting from face-to-face instruction without the enhancement a waste of time and money because the addition does not improve student outcomes.

A second salient feature of online learning practices is the type of pedagogical approach used. Different online pedagogical approaches promote different learning experiences by varying the source of the learning content and the nature of the learner's activity (Galvis, McIntyre, & Hsi, 2006). In traditional didactic or *expository* approaches, content is instructor- or computer directed and typically presented in the form of text, lecture, or instructor-directed discussion. Expository approaches are often contrasted with *active learning*, in which the student engages in exercises and typically proceeds at his or her own pace. Another category of pedagogical approach stresses collaborative or *interactive learning* activity, in which the nature of the learning content is emergent as learners interact with one another and with a teacher or other knowledge sources. Technologies can support any of these three types of pedagogical approach. Online learning researchers have described a pedagogical shift in online learning environments from transmission of knowledge to support for active and interactive learning as newer technologies have expanded online learning possibilities (Rudestam & Schoenholtz-Read, 2010). Researchers are now using terms such as *distributed learning* (Dede, 2006) or *learning communities* (Riel & Polin, 2004; Schwen & Hara, 2004) to refer to orchestrated mixtures of face-to-face and virtual interactions among a cohort of learners led by one or more instructors, facilitators, or coaches over an extended period (from weeks to years).

Finally, a third characteristic commonly used in the past to categorize online learning activities is the extent to which the activity is *synchronous*, with instruction occurring in real time, whether in a physical or a virtual place, or *asynchronous*, with a time lag between the presentation of instructional stimuli and student responses, allowing communication and collaboration over a period of time from anywhere and anytime (Jonassen, Lee, Yang, & Laffey, 2005). An earlier meta-analysis of distance learning applications (Bernard et al., 2004) reported that asynchronous distance education (which included traditional correspondence courses and online courses) had a small but significant advantage over face-to-face instruction in terms of student learning, whereas synchronous distance education (mostly teleconferencing and satellite-based delivery of classes) had a small but significant negative effect. Current forms of online learning support greater interactivity in both synchronous and asynchronous modes, and both communication strategies have their

adherents (Rudestam & Shoenholtz-Read, 2010). A synchronous activity offers greater spontaneity, making learners feel "in synch" with others, thus theoretically promoting collaboration (Hermann, Rummel, & Spada, 2001; Shotsberger, 1999); however, students may feel hurried to respond or hampered by technology breakdowns. In contrast, asynchronous activity offers greater flexibility to learners because it allows them to respond at their convenience. In addition, some argue that the time lag offered in an asynchronous activity allows for more thoughtful and reflective learner participation (Bhattacharya, 1999; Veerman & Veldhuis-Diermanse, 2001), enabling "richer discussions involving more participants" (Dede, 2000). Some have reasoned further that the asynchronous model has more potential to produce a learner-centered environment by encouraging interpersonal, two-way communications between the instructor and an individual student, as well as among students (Bates, 1997).

The meta-analysis reported here examined the influence of these and other learning practices as well as a variety of conditions and methodological features on the online learning effect size by conducting a series of moderator analyses.

RELATED META-ANALYSES

Gene Glass and his colleague pioneered the development of meta-analysis techniques for the systematic quantitative synthesis of results from a series of studies in the 1970s (M. L. Smith & Glass, 1977). Meta-analysis has been used in a variety of fields to inform policy or the design of new research based on the best available evidence (Borenstein, Hedges, Higgins, & Rothstein, 2009). Meta-analysis makes it possible to synthesize data from multiple studies with different sample sizes by extracting an effect size from, and computing a summary effect for, all studies. Lipsey and Wilson (2001) have articulated the following advantages of meta-analysis: (1) it requires an explicit and systematic process for reviewing existing research, therefore enabling the reader to assess the meta-analyst's assumptions, procedures, evidence, and conclusions; (2) it provides a more differentiated and sophisticated summary of existing research than qualitative summaries or "vote counting" on statistical significance by taking into consideration the strength of evidence from each empirical study; (3) it produces synthesized effect estimates with considerably more statistical power than individual studies and allows an examination of differential effects related to different study features; and (4) it provides an organized way of handling information from a large body of studies.

Several meta-analyses related to distance education have been published (Bernard et al., 2004; Machtmes & Asher, 2000; Zhao et al., 2005). Typically these meta-analyses included studies of older generations of technologies such as audio, video, or satellite transmission. One of the most comprehensive meta-analyses on distance education was conducted by Bernard and his colleagues (Bernard et al., 2004). This study examined 699 independent effect sizes from 232 studies published from 1985 to 2001, comparing distance education with classroom instruction for a variety of learners, from young children to adults, on measures of achievement, attitudes, and course completion. The meta-analysis found an overall effect size close to zero for student achievement ($g+ = 0.01$). As noted, asynchronous distance education had a small but significant positive effect ($g+ = 0.05$) on student achievement, whereas synchronous distance education had a small but significant negative effect ($g+ = -0.10$). Bernard et al. found also that a substantial proportion of the variability in effect sizes for student achievement and attitude outcomes was accounted for by the studies' research methodology.

Another meta-analysis of distance education by Zhao and his colleagues (2005) examined 98 effect sizes from 51 studies published from 1996 to 2002. Like Bernard et al.'s study, this meta-analysis focused on distance education courses delivered via multiple generations of technology for a wide variety of learners and found an overall effect size near zero ($d = +0.10$). Subsequent moderator analyses found that studies of blended approaches in which 60%–80% of learning was mediated via technology found significantly more positive effects relative to face-to-face instruction than pure distance learning studies did. The difference between blended learning and classroom instruction was much larger than that between distance education that was almost entirely mediated by technology and classroom instruction. Like the Bernard et al. meta-analysis, that by Zhao et al. included a wide range of outcomes (e.g., achievement, beliefs and attitudes, satisfaction, student dropout rate). Zhao et al. averaged the different kinds of outcomes used in a study to compute an overall effect size for the meta-analysis. This practice is problematic because factors, particularly course features and implementation practices, that enhance one type of student outcome (e.g., student retention) may be quite different from those that enhance another type of outcome (e.g., student achievement) and may even work to the detriment of that other outcome. When mixing studies with different kinds of outcomes, such trade-offs may obscure the relationships between practices and learning.

Some meta-analytic studies have focused on the efficacy of the new generation of distance education courses offered over the Internet for

particular learner populations. Sitzmann et al. (2006), for example, examined 96 studies published from 1996 to 2005 that compared web-based training with face-to-face training for job-related knowledge or skills. The authors found that in general, web-based training was slightly more effective than face-to-face training for acquiring declarative knowledge ("knowing that"), but not for procedural knowledge ("knowing how"). Complicating interpretation of this finding was the fact that Sitzmann et al. found a positive effect of Internet-based training on declarative knowledge in quasi-experimental studies (d = +0.18), but a negative effect favoring face-to-face training in experimental studies with random assignment (d = -0.26). This pattern of findings underscores the need to pay attention to elements of the design of the studies included in a meta-analysis.

Another meta-analysis of online learning by Cavanaugh, Gillan, Kromerey, Hess, and Blomeyer (2004) focused on Internet-based distance education programs for K-12 students. The researchers combined 116 outcomes from 14 studies published between 1999 and 2004 to compute an overall weighted effect, which was not statistically different from zero (g = -0.03). Subsequent investigation of moderator variables found no significant factors affecting student achievement. This meta-analysis used multiple outcomes from the same study, ignoring the fact that the different outcomes from the same student would not be independent of each other. Additionally, the approach used by Cavanaugh et al. assigns more weight to studies with more outcomes than to studies with fewer outcomes (Borenstein et al., 2009).

In summary, although some meta-analytic studies have investigated the outcomes of distance education for a wide range of learners (Bernard et al., 2004; Zhao et al., 2005), none of the large-scale meta-analyses applying methodological quality standards in the selection of studies isolated learning outcomes with Internet-based learning environments from other types of outcomes and from older distance education technologies. Advances in Internet-based learning tools and their increased popularity across different learning contexts warrants a rigorous meta-analysis of student learning outcomes with online learning. Past meta-analyses typically included studies with weak research designs (i.e., quasi-experimental studies without statistical control for preexisting differences), thereby summarizing findings that are themselves subject to threats to internal validity. The finding in several meta-analyses that the size of a study's effect is related inversely to research design quality (Bernard et al., 2004; Pearson, Ferdig, Blomeyer, & Moran, 2005; Sitzmann et al., 2006) implies the need for computing an overall online learning effect with data drawn exclusively from studies with acceptably rigorous research designs.

PURPOSE OF THE PRESENT STUDY

This meta-analysis was conducted to examine the effectiveness of both purely online and blended versions of online learning as compared with traditional face-to-face learning. Our approach differs from prior meta-analyses of distance learning in several important respects:

- Only studies of web-based learning have been included (i.e., eliminating studies of video- and audio-based telecourses or stand-alone, computer-based instruction).
- Only studies with random-assignment or controlled quasi-experimental designs have been included to draw on the best available evidence.
- All effects have been based on objective and direct measures of learning (i.e., discarding effects for student or teacher perceptions of learning, their satisfaction, retention, attendance, etc.).

In addition to examining the learning effects of the two forms of online learning—namely, purely online learning and blended learning—relative to face-to-face learning, this meta-analysis investigated a series of conditions and practices that may be associated with differences in the effectiveness of online instruction. Conditions investigated include the year in which the intervention took place, the learners' demographic characteristics, and the teacher's or instructor's training. In contrast to conditions, which are not subject to the practitioner's control, *practices* concern how online learning is implemented (e.g., whether online students had the opportunity to interact with an instructor). The meta-analysis sought to examine practices such as the duration of the intervention, provision of synchronous computer-mediated communication, and the incorporation of learner feedback.

Four research questions guided the study design and analysis:

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 (i.e., blended instruction) enhance learning?
3. What practices are associated with more effective online learning?
4. What conditions influence the effectiveness of online learning?

METHOD

LITERATURE SEARCH

Relevant studies were located through a comprehensive search of publicly available literature published from 1996 through July 2008. We chose 1996 as a starting point for the literature search because web-based learning resources and tools became widely available around that time. The following data sources and search tools were used: (1) electronic research databases, including ERIC, PsycINFO, PubMed, ABI/INFORM, and UMI ProQuest Digital Dissertations. Search strategies were adapted to fit the tool used, but all searches were conducted with combinations of two types of search terms, one an education or training term (e.g., *distance education, e-learning, online learning, distributed learning*), and the other a study design term (e.g., *control group, comparison group, treatment group, experimental*); (2) articles cited in recent meta-analyses and narrative syntheses of research on distance learning (Bernard et al., 2004; Cavanaugh et al., 2004; Childs, 2001; Sitzmann et al., 2006; Tallent-Runnels et al., 2006; WestEd with Edvance Research, 2008; Whitehouse, Breit, McCloskey, Ketelhut, & Dede 2006; Wisher & Olson, 2003; Zhao et al., 2005; Zirkle 2003); (3) articles published since 2005 in the following key journals: *American Journal of Distance Education, Journal of Distance Education* (Canada), *Distance Education* (Australia), *International Review of Research in Distance and Open Education, Journal of Asynchronous Learning Networks, Journal of Technology and Teacher Education,* and *Career and Technical Education Research*; and (4) the Google Scholar search engine with a subset of the search terms used in the electronic research databases.[2]

SCREENING PROCESS: INCLUSION AND EXCLUSION CRITERIA

Screening of the research studies obtained through the data sources described earlier was carried out in two stages: abstract screening of the initial electronic database searches and full-text screening of studies that passed the abstract screen. The intent of the two-stage approach was to gain efficiency without risking exclusion of potentially relevant, high-quality studies of online learning effects.

The initial electronic database searches yielded 1,132 articles (including duplicates of the same article returned by different databases). Citation information and abstracts of these studies were examined to ascertain whether they met the following three initial inclusion criteria: (1) the study addresses online learning as this study defines it; (2) the

study appears to use a controlled design (experimental/controlled quasi-experimental design); and (3) the study reports data on student achievement or another learning outcome. At this early stage, analysts gave studies "the benefit of the doubt," retaining those that were not clearly outside the inclusion criteria on the basis of their citations and abstracts.

As a result of this screening, 316 articles were retained, and 816 articles were excluded. During this initial screen, 45% of the articles were excluded primarily because they did not have a controlled design; 26% of articles were eliminated because they did not report learning outcomes for treatment and control groups; and 23% were eliminated because their intervention did not qualify as online learning, given the definition used for this meta-analysis and review. The remaining 6% of the articles posed other difficulties, such as being written in a language other than English. From the other data sources (i.e., references in earlier reviews, manual review of key journals, recommendation from a study advisor, and Google Scholar searches), an additional 186 articles were retrieved, yielding a total of 502 articles that were subjected to a full-text screening for possible inclusion in the analysis.

A set of full-text screening criteria was applied against each of the 502 articles. The screening criteria included both topical relevance and study methodology. A study had to meet *content relevance* criteria to be included in the meta-analysis. Qualifying studies had to:

- *Involve learning that took place over the Internet.* The use of the Internet had to be a substantial part of the intervention. Studies in which the Internet was only an incidental component of the intervention were excluded. In operational terms, to qualify as *online learning*, a study treatment needed to provide at least a quarter of the instruction/learning of the content assessed by the study's learning measure by means of the Internet.
- *Contrast conditions that varied in terms of use of online learning.* Student outcomes for classroom-based instruction had to be compared against a condition falling into at least one of two study categories: (1) purely online learning compared with offline/face-to-face learning, or (2) a combination of online plus offline/face-to-face learning (i.e., blended learning) compared with offline/face-to-face learning alone.
- *Describe an intervention study that had been completed.* Descriptions of study designs, evaluation plans, or theoretical frameworks were excluded. The length of the intervention/treatment could vary from a few hours to a quarter, semester, year, or longer.
- *Report a learning outcome that was measured for both treatment and control*

groups. A learning outcome needed to be measured in the same way across study conditions. A study was excluded if it explicitly indicated that different examinations were used for the treatment and control groups. The measure had to be objective and direct; learner or teacher/instructor reports of learning were not considered direct measures. Examples of acceptable learning outcome measures included scores on standardized tests, scores on researcher-created assessments, grades/scores on teacher-created assessments (e.g., assignments, midterm/final exams), and grades or grade point averages. Examples of learning outcome measures for teacher learners (in addition to those accepted as student outcomes) included assessments of content knowledge, analysis of lesson plans or other materials related to the intervention, observation (or logs) of class activities, analysis of portfolios, or supervisor's rating of job performance. Studies that used only nonlearning outcome measures (e.g., attitude, retention, attendance, level of learner/instructor satisfaction) were excluded.

Studies also had to meet basic *methodology* criteria to be included. Qualifying studies had to:

• *Use a controlled design* (experimental or quasi-experimental). Contemporary standards for meta-analyses focusing on the effectiveness of interventions call for restricting the corpus of studies to true experiments and high-quality quasi-experiments (Bethel & Bernard, 2010). Design studies, exploratory studies, or case studies that did not use a controlled research design were excluded. For quasi-experimental designs, the analysis of the effects of the intervention had to include statistical controls for possible differences between the treatment and control groups in terms of prior achievement.

To ensure the reliability of the full-text screening, nine analysts were trained on the full-text screening criteria and practiced their application with several training articles. After the training, analysts read full-text articles independently but were asked to bring up all borderline cases for discussion and resolution either at project meetings or through consultation with task leaders. To prevent studies from being mistakenly screened out, two analysts coded studies on features that were deemed to require significant degrees of inference. These features consisted of the following: (1) failure to have students use the Internet for a substantial portion of the time that they spent learning the content assessed by the

study's learning measure, and (2) lack of statistical control for prior abilities in quasi-experiments.

From the 502 articles, analysts identified 522 independent studies (some articles reported more than one study). When the same study was reported in different publication formats (e.g., conference paper and journal article), only the more formal journal article was retained for the analysis. Of the 522 studies, 176 met all the criteria of the full-text screening process. Table 1 shows the bases for exclusion for the 346 studies that did not meet all the criteria.

Table 1. Bases for Excluding Studies During the Full-Text Screening Process

Primary Reason for Exclusion	Number Excluded	Percentage Excluded
Did not use statistical control	137	39
Was not online as defined in this review	90	26
Did not analyze learning outcomes	52	15
Did not have a comparison group that received a comparable treatment	22	7
Did not fit into any of the three study categories	39	11
Excluded for other reasons[a]	6	2

[a]Other reasons for exclusion included: (1) did not provide enough information, (2) was written in a language other than English, and (3) used different learning outcome measures for the treatment and control groups.

EFFECT SIZE EXTRACTION

Of the 176 independent studies, 99 had at least one contrast between purely online learning and face-to-face/offline learning or between blended learning and face-to-face/offline learning. These studies were subjected to quantitative analysis to extract effect sizes. Two senior analysts examined the 99 studies to extract the information needed for calculating or estimating an effect size. To avoid eliminating some articles that might actually have had the needed statistical data, a second analyst reviewed those cases considered for elimination on the grounds of inadequate data.

Following the guidelines from the What Works Clearinghouse (2007) and Lipsey and Wilson (2001), numerical and statistical data contained in the studies were extracted so that Comprehensive Meta-Analysis software (Biostat Solutions, 2006) could be used to calculate effect sizes (g).

The precision of each effect estimate was determined by using the estimated standard error of the mean to calculate the 95% confidence interval for each effect.

During the data extraction phase, it became apparent that one set of studies rarely provided sufficient data for Comprehensive Meta-Analysis calculation of an effect size. Quasi-experimental studies that used hierarchical linear modeling or analysis of covariance with adjustment for pretests and other learner characteristics through covariates typically did not report some of the data elements needed to compute an effect size. For studies using hierarchical linear modeling to analyze impacts, typically the regression coefficient on the treatment status variable (treatment or control), its standard error, and a p value and sample sizes for the two groups were reported. For analyses of covariance, typically the adjusted means and F statistic were reported, along with group sample sizes. In almost all cases, the unadjusted standard deviations for the two groups were not reported and could not be computed because the pretest-posttest correlation was not provided. To avoid eliminating all these studies (which included some of the largest and most recent investigations), analysts used a conservative estimate of the pretest-posttest correlation ($r = .70$) in order to estimate an effect size for those studies in which the pretest was the same measure as the posttest, and a pretest-posttest correlation of $r = .50$ for studies in which different measures were used at pretest and posttest. These effect sizes were flagged in the coding as "estimated effect sizes," as were effect sizes computed from t tests, F tests, and p levels. In extracting effect size data, analysts followed a set of rules:

- The unit of analysis was the independent contrast between the online condition and the face-to-face condition or between the blended condition and the face-to-face condition. Some studies reported more than one contrast, either by reporting more than one experiment or by having multiple treatment conditions (e.g., online vs. blended vs. face-to-face) in a single experiment.
- When there were multiple treatment groups or multiple control groups and the nature of the instruction in the groups did not differ considerably (e.g., two treatment groups both falling into the "blended" instruction category), the weighted mean of the groups and pooled standard deviation were used.
- When there were multiple treatment groups or multiple control groups and the nature of the instruction in the groups differed considerably (e.g., one treatment was purely online whereas the other treatment was blended instruction, both compared against the face-

to-face condition), analysts treated them as independent contrasts.

- In general, one learning outcome finding was extracted from each study. When multiple learning outcome data were reported (e.g., assignments, midterm and final examinations, grade point averages, grade distributions), the outcome that could be expected to be more stable and more closely aligned to the instruction was extracted (e.g., final examination scores instead of quizzes). However, in some studies, no learning outcome had obvious superiority over the others. In such cases, analysts extracted multiple contrasts from the study and calculated the weighted average of the multiple outcome scores if the outcome measures were similar (e.g., two tests of similar length and content) but retained both measures if they addressed different kinds of learning (for example, a multiple-choice knowledge test and a performance-based assessment of strategic and problem-solving skills applied to ill-structured problems).

- Learning outcome findings were extracted at the individual level. Analysts did not extract group-level learning outcomes (e.g., scores for a group product). Too few group products were included in the studies to support analyses of this variable.

The review of the 99 studies to obtain the data for calculating effect size produced 50 independent effect sizes (27 for purely online vs. face-to-face and 23 for blended vs. face-to-face) from 45 studies. Fifty-four studies did not report sufficient data to support calculating an effect size.

CODING OF STUDY FEATURES

All studies that provided enough data to compute an effect size were coded for practices, conditions, and features of study methodology. Building on the project's conceptual framework (Figure 2) and the coding schemes used in several earlier meta-analyses (Bernard et al., 2004; Sitzmann et al., 2006), a coding structure was developed and pilot-tested with several studies. The top-level coding structure, incorporating refinements made after pilot testing, is shown in Table 2.

To determine interrater reliability, two researchers coded 20% of the studies, achieving an interrater reliability of 86% across those studies. Analysis of coder disagreements resulted in the refinement of some definitions and decision rules for some codes; other codes that required information that articles did not provide or that proved difficult to code reliably were eliminated (e.g., whether or not the online instructor had been trained in this method of instruction). A single researcher coded the remaining studies.

Table 2. Top-Level Coding Structure for the Meta-Analysis

Study Condition Codes

- Author
- Learner age
- Learner incentive for involvement in the study
- Learner type
- Learning setting
- Subject matter
- Type of publication
- Year of publication

Online Learning Practice Codes

- Dominant approach to learner control
- Media features
- Nature of knowledge assessed
- Nature of outcome measure
- Opportunity for *asynchronous* computer-mediated communication with the instructor
- Opportunity for *asynchronous* computer-mediated communication with peers
- Opportunity for face-to-face contact with the instructor
- Opportunity for face-to-face contact with peers
- Opportunity for *synchronous* computer-mediated communication with the instructor
- Opportunity for *synchronous* computer-mediated communication with peers
- Opportunity for feedback
- Opportunity for practice
- Pedagogical approach
- Treatment duration
- Use of problem-based or project-based learning
- Whether the instructor was trained in online teaching

Study Method Codes

- Attrition equivalence
- Contamination
- Curriculum material/instruction equivalence
- Equivalence of prior knowledge/pretest scores
- Instructor equivalence
- Sample size for unit of assignment
- Student equivalence
- Study design
- Time-on-task equivalence
- Unit of assignment to conditions
- Whether equivalence of groups at preintervention was described

DATA ANALYSIS

Before combining effects from multiple contrasts, effect sizes were weighted to avoid undue influence of studies with small sample sizes (Hedges & Olkin, 1985). For the total set of 50 contrasts and for each subset of contrasts being investigated, a weighted mean effect size (Hedges' $g+$) was computed by weighting the effect size for each study contrast by the inverse of its variance. The precision of each mean effect estimate was determined by using the estimated standard error of the mean to calculate the 95% confidence interval. Using a fixed-effects model, the heterogeneity of the effect size distribution (the Q-statistic) was computed to indicate the extent to which variation in effect sizes was not explained by sampling error alone.

Next, a series of post-hoc subgroup and moderator variable analyses was conducted using the Comprehensive Meta-Analysis software. A mixed-effects model was used for these analyses to model within-group variation. In comparison with a fixed-effects model, the mixed-effects model reduces the likelihood of Type I errors by adding a random constant to the standard errors, but it does so at the cost of increasing the likelihood of Type II errors (incorrectly accepting the null hypothesis).

A between-group heterogeneity statistic (QBetween) was computed to test for statistical differences in the weighted mean effect sizes for various subsets of the effects (e.g., studies using blended as opposed to purely online learning for the treatment group).

FINDINGS

NATURE OF THE STUDIES IN THE META-ANALYSIS

As noted, 50 independent effect sizes could be abstracted from the study corpus of 45 studies. The number of learners in the studies included in the meta-analysis ranged from 16 to 1,857, but most of the studies were modest in scope. Although large-scale applications of online learning have emerged, only five studies in the meta-analysis corpus included more than 400 learners. The types of learners in the studies in the meta-analysis were about evenly split between students in college or earlier years of education and learners in graduate programs or professional training. The average learner age in a study ranged from 13 to 44 years. Nearly all the studies involved formal instruction, with the most common subject matter being medicine or health care. Other content types included computer science, teacher education, social science, mathematics, languages, science, and business. Roughly half of the learners were

taking the instruction for credit or as an academic requirement. Of the 48 contrasts for which the study indicated the length of instruction, 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 predominantly instructor directed (8 contrasts) than they were to be predominantly student directed, independent learning (17 contrasts), or interactive and collaborative in nature (22 contrasts). Online learners typically had opportunities to practice skills or test their knowledge (41 effects were from studies reporting such opportunities). Opportunities for learners to receive feedback were less common; however, it was reported in the studies associated with 23 effects. The opportunity for online learners to have face-to-face contact with the instructor during the time frame of the course was present in the case of 21 out of 50 effects.

The details of instructional media and communication options available to online learners were absent in many of the study narratives. Among the 50 contrasts, analysts could document the presence of one-way video or audio in the online condition for 14 effects. Similarly, 16 contrasts involved online conditions that allowed students to communicate with the instructor with asynchronous communication only; 8 allowed both asynchronous and synchronous online communication; and 26 contrasts came from studies that did not document the types of online communication provided between the instructor and learners.

Among the 50 individual contrasts between online and face-to-face instruction, 11 were significantly positive, favoring the online or blended learning condition. Three significant negative effects favored traditional face-to-face instruction. That multiple comparisons were conducted should be kept in mind when interpreting this pattern of findings.

Figure 3 illustrates the 50 effect sizes derived from the 45 articles. Some references appear twice in Figure 3 because multiple effect sizes were extracted from the same article. Davis, Odell, Abbitt, and Amos (1999) and Caldwell (2006) each included two contrasts—purely online versus face-to-face and blended versus face-to-face. Rockman et al. (2007) and Schilling, Wiecha, Polineni, and Khalil (2006) reported findings for two distinct learning measures. M. Long and Jennings (2005) reported findings from two distinct experiments, a wave 1 in which teachers were implementing online learning for the first time and a wave 2 in which teachers implemented online learning a second time with new groups of students.

Tables 3 and 4 present the effect sizes for purely online versus face-to-face and blended versus face-to-face studies, respectively, along with standard errors, statistical significance, and the 95% confidence interval.

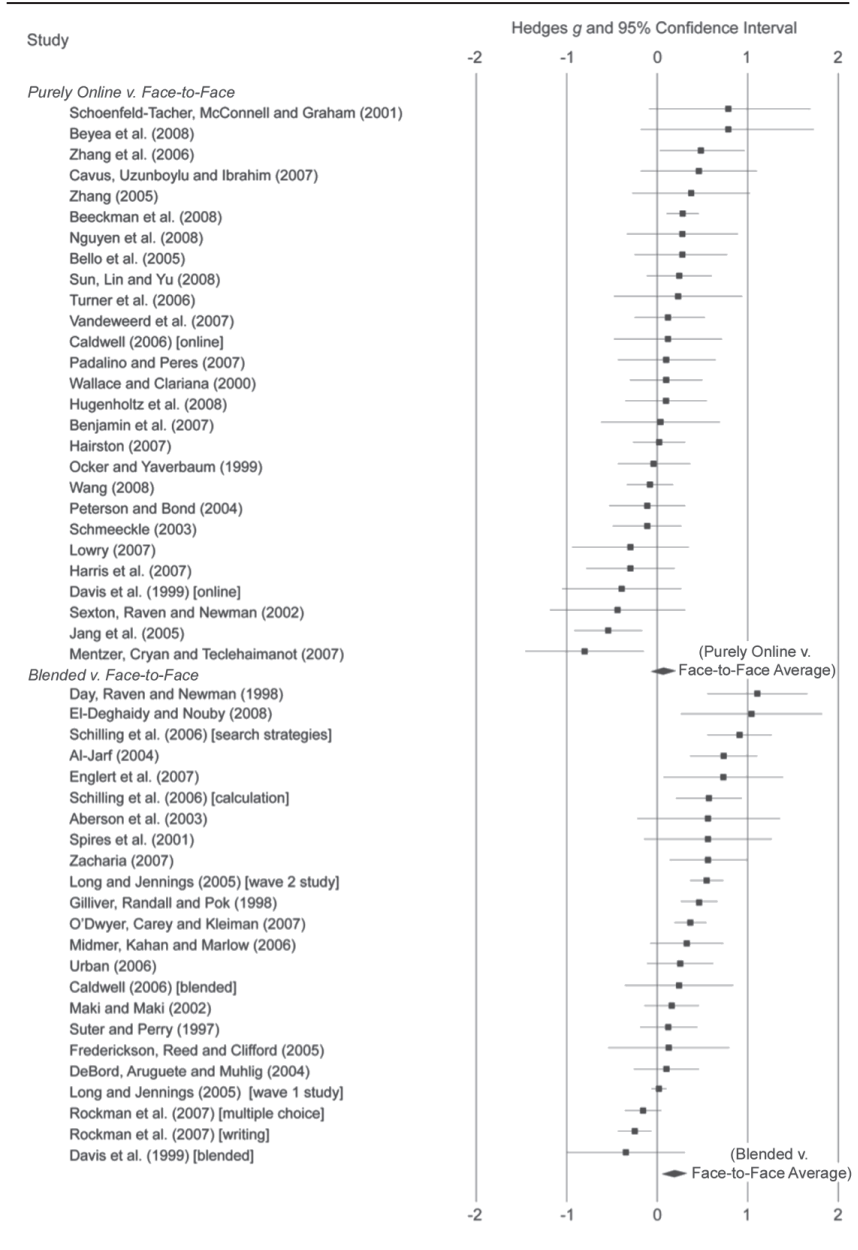**Figure 3. Effect sizes for contrasts in the meta-analysis**

Purely Online v. Face-to-Face
Schoenfeld-Tacher, McConnell and Graham (2001)
Beyea et al. (2008)
Zhang et al. (2006)
Cavus, Uzunboylu and Ibrahim (2007)
Zhang (2005)
Beeckman et al. (2008)
Nguyen et al. (2008)
Bello et al. (2005)
Sun, Lin and Yu (2008)
Turner et al. (2006)
Vandeweerd et al. (2007)
Caldwell (2006) [online]
Padalino and Peres (2007)
Wallace and Clariana (2000)
Hugenholtz et al. (2008)
Benjamin et al. (2007)
Hairston (2007)
Ocker and Yaverbaum (1999)
Wang (2008)
Peterson and Bond (2004)
Schmeeckle (2003)
Lowry (2007)
Harris et al. (2007)
Davis et al. (1999) [online]
Sexton, Raven and Newman (2002)
Jang et al. (2005)
Mentzer, Cryan and Teclehaimanot (2007)

Blended v. Face-to-Face
Day, Raven and Newman (1998)
El-Deghaidy and Nouby (2008)
Schilling et al. (2006) [search strategies]
Al-Jarf (2004)
Englert et al. (2007)
Schilling et al. (2006) [calculation]
Aberson et al. (2003)
Spires et al. (2001)
Zacharia (2007)
Long and Jennings (2005) [wave 2 study]
Gilliver, Randall and Pok (1998)
O'Dwyer, Carey and Kleiman (2007)
Midmer, Kahan and Marlow (2006)
Urban (2006)
Caldwell (2006) [blended]
Maki and Maki (2002)
Suter and Perry (1997)
Frederickson, Reed and Clifford (2005)
DeBord, Aruguete and Muhlig (2004)
Long and Jennings (2005) [wave 1 study]
Rockman et al. (2007) [multiple choice]
Rockman et al. (2007) [writing]
Davis et al. (1999) [blended]

Table 3. Purely Online Versus Face-to-Face Studies Included in the Meta-Analysis

Authors	Title	Effect Size		95-Percent Confidence Interval		Test of Null Hypothesis (2-tail)	Retention Rate (percentage)		Number of Units Assigned[a]
		g	SE	Lower Limit	Upper Limit	Z Value	Online	Face-to-Face	
Beeckman et al. (2008)	Pressure ulcers: E-learning to improve classification by nurses and nursing students	+0.294	0.097	0.104	0.484	3.03**	Unknown	Unknown	426 participants
Bello et al. (2005)	Online vs. live methods for teaching difficult airway management to anesthesiology residents	+0.278	0.265	-0.241	0.797	1.05	100	100	56 participants
Benjamin et al. (2008)	A randomized controlled trial comparing Web to in-person training for child care health consultants	+0.046	0.340	-0.620	0.713	0.14	Unknown	Unknown	23 participants
Beyea et al. (2008)	Evaluation of a particle repositioning maneuver web-based teaching module	+0.790	0.493	-0.176	1.756	1.60	Unknown	Unknown	17–20 participants[b]
Caldwell (2006)	A comparative study of traditional, web-based and online instructional modalities in a computer programming course	+0.132	0.310	-0.476	0.740	0.43	100	100	60 students
Cavus, Uzonboylu, and Ibrahim (2007)	Assessing the success rate of students using a learning management system together with a collaborative tool in web-based teaching of programming languages	+0.466	0.335	-0.190	1.122	1.39	Unknown	Unknown	54 students
Davis et al. (1999)	Developing online courses: A comparison of web-based instruction with traditional instruction	-0.379	0.339	-1.042	0.285	-1.12	Unknown	Unknown	2 courses/classrooms
Hairston (2007)	Employees' attitudes toward e-learning: Implications for policy in industry environments	+0.028	0.155	-0.275	0.331	0.18	70	58.33	168 participants
Harris et al. (2008)	Educating generalist physicians about chronic pain: Live experts and online education can provide durable benefits	-0.285	0.252	-0.779	0.209	-1.13	84.21	94.44	62 participants

Study	Title									
Hugenholtz et al. (2008)	Effectiveness of e-learning in continuing medical education for occupational physicians	+0.106	0.233	-0.351	0.564	0.46	Unknown	Unknown	72 participants
Jang et al. (2005)	Effects of a Web-based teaching method on undergraduate nursing students' learning of electrocardiography	-0.530	0.197	-0.917	-0.143	-2.69**	85.71	87.93	105 students
Lowry (2007)	Effects of online versus face-to-face professional development with a team-based learning community approach on teachers' application of a new instructional practice	-0.281	0.335	-0.937	0.370	-0.84	80	93.55	53 students
Mentzer, Cryan and Teclehaimanot (2007)	A comparison of face-to-face and web-based classrooms	-0.796	0.339	-1.460	-0.131	-2.35*	Unknown	Unknown	36 students
Nguyen et al. (2008)	Randomized controlled trial of an Internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: Pilot study	+0.292	0.316	-0.327	0.910	0.93	Unknown	Unknown	39 participants
Ocker and Yaverbaum (1999)	Asynchronous computer-mediated communication versus face-to-face collaboration: Results on student learning, quality and satisfaction	-0.030	0.214	-0.449	0.389	-0.14	Unknown	Unknown	43 students
Padalino and Peres (2007)	E-learning: A comparative study for knowledge apprehension among nurses	0.115	0.281	-0.437	0.666	0.41	Unknown	Unknown	49 participants
Peterson and Bond (2004)	Online compared to face-to-face teacher preparation for learning standards-based planning skills	-0.100	0.214	-0.520	0.320	-0.47	Unknown	Unknown	4 sections
Schmeeckle (2003)	Online training: An evaluation of the effectiveness and efficiency of training law enforcement personnel over the Internet	-0.106	0.198	-0.494	0.282	-0.53	Unknown	Unknown	101 students

Table 3. Purely Online Versus Face-to-Face Studies Included in the Meta-Analysis (continued)

Study	Description								
Schoenfeld-Tacher, McConnell and Graham (2001)	Do no harm: A comparison of the effects of online vs. traditional delivery media on a science course	+0.800	0.459	-0.100	1.700	1.74	100	99.94	Unknown
Sexton et al. (2002)	A comparison of traditional and World Wide Web methodologies, computer anxiety, and higher order thinking skills in the inservice training of Mississippi 4-H extension agents	-0.422	0.385	-1.177	0.332	-1.10	Unknown	Unknown	26 students
Turner et al. (2006)	Web-based learning versus standardized patients for teaching clinical diagnosis: A randomized, controlled, crossover trial	+0.242	0.367	-0.477	0.960	0.66	Unknown	Unknown	4 classrooms
Vandeweerd et al. (2007)	Teaching veterinary radiography by e-learning versus structured tutorial: A randomized, single-blinded controlled trial	+0.144	0.207	-0.262	0.550	0.70	Unknown	Unknown	30 students
Wallace and Clariana (2000)	Achievement predictors for a computer-applications module delivered online	+0.109	0.206	-0.295	0.513	0.53	Unknown	Unknown	92 students
Wang (2008)	Developing and evaluating an interactive multimedia instructional tool: Learning outcomes and user experiences of optometry students	-0.071	0.136	-0.338	0.195	-0.53	Unknown	Unknown	4 sections
Zhang (2005)	Interactive multimedia-based e-learning: A study of effectiveness.	+0.381	0.339	-0.283	1.045	1.12	Unknown	Unknown	51 students
Zhang et al. (2006)	Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness	+0.499	0.244	0.022	0.977	2.05*	Unknown	Unknown	69 students

[a] The number given represents the assigned units at study conclusion. It excludes units that attrited.

[b] Two outcome measures were used to compute one effect size. The first outcome measure was completed by 17 participants, and the second outcome measure was completed by 20 participants.

[c] This study is a crossover study. The number of units represents those assigned to treatment and control conditions in the first round.

*$p < .05$. **$p < .01$. SE = standard error.

Table 4. Blended Versus Face-to-Face Studies Included in the Meta-Analysis

Authors	Title	Effect Size		95-Percent Confidence Interval		Test of Null Hypothesis (2-tail)	Retention Rate (percentage)		Number of Units Assigned[a]
		g	SE	Lower Limit	Upper Limit	ZValue	Online	Face-to-Face	
Aberson, Berger, and Romero (2003)	Evaluation of an interactive tutorial for teaching hypothesis testing concepts	+0.580	0.404	-0.212	1.372	1.44	Unknown	.75	2 sections
Al-Jarf (2004)	The effects of Web-based learning on struggling EFL college writers	+0.740	0.194	0.360	1.120	3.82***	Unknown	Unknown	113 students
Caldwell (2006)	A comparative study of three instructional modalities in a computer programming course	+0.251	0.311	-0.359	0.861	0.81	100	100	60 students
Davis et al. (1999)	Developing online courses: A comparison of Web-based instruction with traditional instruction	-0.335	0.338	-0.997	0.327	-0.99	Unknown	Unknown	2 courses/classrooms
Day, Raven, and Newman (1998)	The effects of World Wide Web instruction and traditional instruction and learning styles on achievement and changes in student attitudes in a technical writing in agricommunication course	+1.113	0.289	0.546	1.679	3.85***	89.66	96.55	2 sections
DeBord, Aruguete and Muhlig (2004)	Are computer-assisted teaching methods effective?	+0.110	0.188	-0.259	0.479	0.69	Unknown	Unknown	112 students
El-Deghaidy and Nouby (2008)	Effectiveness of a blended e-learning cooperative approach in an Egyptian teacher education programme	+1.049	0.406	0.253	1.845	2.58**	Unknown	Unknown	26 students
Englert et al. (2007)	Scaffolding the writing of students with disabilities through procedural facilitation using an Internet-based technology to improve performance	+0.740	0.345	0.064	1.416	2.15*	Unknown	Unknown	6 classrooms from 5 urban schools

Table 4.

Table 4. Blended Versus Face-to-Face Studies Included in the Meta-Analysis (continued)

Study	Title								
Frederickson, Reed and Clifford (2005)	Evaluating Web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives	+0.138	0.345	-0.539	0.814	0.40	Unknown	Unknown	2 sections
Gilliver, Randall, and Pok (1998)	Learning in cyberspace: Shaping the future	+0.477	0.111	0.260	0.693	4.31***	Unknown	Unknown	24 classes
Long and Jennings (2005) [Wave 1] [c]	"Does it work?": The effect of technology and professional development on student achievement	+0.025	0.046	-0.066	0.116	0.53	Unknown	Unknown	9 schools
Long and Jennings (2005) [Wave 2] [c]	"Does it work?": The effect of technology and professional development on student achievement	+0.554	0.098	0.362	0.747	5.65***	Unknown	Unknown	6 teachers
Maki and Maki (2002)	Multimedia comprehension skill predicts differential outcomes of Web-based and lecture courses	+0.171	0.160	-0.144	0.485	1.06	91.01	88.10	155 students
Midmer, Kahan, and Marlow (2006)	Effects of a distance learning program on physicians' opioid- and benzodiazepine-prescribing skills	+0.332	0.213	-0.085	0.750	1.56m	Unknown	Unknown	88 students
O'Dwyer, Carey, and Kleiman (2007)	A study of the effectiveness of the Louisiana algebra I online course	+0.373	0.094	0.190	0.557	3.99***	88.51	64.4	Unknown [b]
Rockman et al. (2007) [Writing] [c]	ED PACE final report	-0.239	0.102	-0.438	-0.039	-2.34*	Unknown	Unknown	28 classrooms
Rockman et al. (2007) [Multiple-choice test] [c]	ED PACE final report	-0.146	0.102	-0.345	0.054	-1.43	Unknown	Unknown	28 classrooms
Schilling et al. (2006) [Search strategies] [c]	An interactive Web-based curriculum on evidence-based medicine: Design and effectiveness	+0.585	0.188	0.216	0.953	3.11**	68.66	59.62	Unknown

27

Study	Title								
Schilling et al. (2006) [Quality of care calculation] c	An interactive Web-based curriculum on evidence-based medicine: Design and effectiveness	+0.926	0.183	0.567	1.285	5.05***	66.42	86.54	Unknown
Spires et al. (2001)	Exploring the academic self within an electronic mail environment	+0.571	0.357	-0.130	1.271	1.60	Unknown	100.00	31 students
Suter and Perry (1997)	Evaluation by electronic mail	+0.140	0.167	-0.188	0.468	0.84	Unknown	Unknown	Unknown
Urban (2006)	The effects of using computer-based distance education for supplemental instruction compared to traditional tutorial sessions to enhance learning for students at-risk for academic difficulties	+0.264	0.192	-0.112	0.639	1.37	96.86	73.85	110 students
Zacharia (2007)	Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits	+0.570	0.216	0.147	0.993	2.64**	100	95.56	88 students

a This number represents the assigned units at study conclusion. It excludes units that attrited.

b The study involved 18 online classrooms from six districts and two private schools; the same six districts were asked to identify comparable face-to-face classrooms, but the study does not report how many of those classrooms participated.

c Two independent contrasts were contained in this article, which therefore appears twice in the table.

*$p < .05$. **$p < .01$. ***$p < .001$. SE = standard error.

MAIN EFFECTS

The overall finding of the meta-analysis is that online learning (the combination of studies of purely online and of blended learning) on average produces stronger student learning outcomes than learning solely through face-to-face instruction. The mean effect size for all 50 contrasts was +0.20, $p < .001$.

Next, separate mean effect sizes were computed for purely online versus face-to-face and blended versus face-to-face contrasts. The mean effect size for the 27 purely online versus face-to-face contrasts was not significantly different from 0 ($g+ = +0.05$, $p = .46$). The mean effect size for the 23 blended versus face-to-face contrasts was significantly different from 0 ($g+ = +0.35$, $p < .0001$).

A test of the difference between the purely online versus face-to-face studies and the blended versus face-to-face studies found that the mean effect size was larger for contrasts pitting blended learning against face-to-face instruction than for those of purely online versus face-to-face instruction ($Q = 8.37$, $p < .01$). Thus, studies of blended instruction found a larger advantage relative to face-to-face instruction than did studies of purely online learning.

TEST FOR HOMOGENEITY

Analysts used the entire corpus of 50 effects to explore the influence of possible moderator variables. The individual effect size estimates included in this meta-analysis ranged from a low of −0.80 (higher performance in the face-to-face condition) to a high of +1.11 (favoring online instruction). A test for homogeneity of effect size found significant differences across studies ($Q = 168.86$, $p < .0001$). This significant heterogeneity in effect sizes justifies the investigation of the variables that may have influenced the differing effect sizes.

ANALYSES OF MODERATOR VARIABLES

The study's conceptual framework identifies practice and condition variables that might be expected to correlate with the effectiveness of online learning as well as study method variables, which often correlate with effect size. Typically, more poorly controlled studies show larger effects. Each study in the meta-analysis was coded for these three types of vari-

ables—practices, conditions, and study methods—using the coding cate-
gories shown in Table 2.

Many of the studies did not provide information about features consid-
ered to be potential moderator variables, a predicament noted in previ-
ous meta-analyses (see Bernard et al., 2004). Many of the reviewed
studies, for example, did not indicate rates of attrition from the contrast-
ing conditions or evidence of contamination between conditions.

For some of the variables, the number of studies providing sufficient
information to support categorization as to whether the feature was pre-
sent was too small to support a meaningful analysis. Analysts identified
those variables for which at least two contrasting subsets of studies, with
each subset containing six or more study effects, could be constructed. In
some cases, this criterion could be met by combining related feature
codes; in a few cases, the inference was made that failure to mention a
particular practice or technology (e.g., one-way video) denoted its
absence. Practice, condition, and method variables for which study sub-
sets met the size criterion were included in the search for moderator vari-
ables.

PRACTICE VARIABLES

Table 5 shows the variation in effectiveness associated with 12 practice
variables. Table 5 and the two data tables that follow show significance
results both for the various subsets of studies considered individually and
for the test of the dimension used to subdivide the study sample (i.e., the
potential moderator variable). For example, in the case of Synchronicity
of Communication With Peers, both the 17 contrasts in which students in
the online condition had only asynchronous communication with peers
and the 6 contrasts in which online students had both synchronous and
asynchronous communication with peers are shown in the table. The two
subsets had mean effect sizes of +0.27 and +0.17, respectively, and only
the former was statistically different from 0. The *Q*-statistic of homogene-
ity tests whether the variability in effect sizes for these contrasts is associ-
ated with the type of peer communication available. The *Q*-statistic for
Synchronicity of Communication With Peers (0.32) is not statistically dif-
ferent from 0, indicating that the addition of synchronous communica-
tion with peers is not a significant moderator of online learning
effectiveness.

Table 5. Tests of Practices as Moderator Variables

Variable	Contrast	Number Studies	Weighted Effect Size	Standard Error	Lower Limit	Upper Limit	*Q*-Statistic
Pedagogy/ learning experience[a]	Instructor-directed (expository)	8	0.386**	0.120	0.150	0.622	
	Independent (active)	17	0.050	0.082	-0.110	0.210	6.19*
	Collaborative (interactive)	22	0.249***	0.075	0.102	0.397	
Computer-mediated communication with instructor[a]	Asynchronous only	16	0.239*	0.108	0.027	0.451	
	Synchronous + Asynchronous	8	0.036	0.151	-0.259	0.331	1.20
Computer-mediated communication with peers[a]	Asynchronous only	17	0.272**	0.091	0.093	0.450	
	Synchronous + Asynchronous	6	0.168	0.158	-0.141	0.478	0.32
Treatment duration[a]	Less than 1 month	19	0.140	0.089	-0.034	0.314	
	More than 1 month	29	0.234***	0.069	0.098	0.370	0.69
Media features[a]	Text-based only	14	0.208	0.111	-0.009	0.425	
	Text + other media	32	0.200**	0.066	0.071	0.329	0.00
Time on task[a]	Online > Face to Face	9	0.451***	0.113	0.229	0.673	
	Same or Face to Face > Online	18	0.183*	0.083	0.020	0.346	3.62
One-way video or audio	Present	14	0.092	0.091	-0.087	0.271	
	Absent/Not reported	36	0.254***	0.062	0.133	0.375	2.15
Computer-based instruction elements	Present	29	0.182**	0.065	0.054	0.311	
	Absent/Not reported	21	0.234**	0.081	0.075	0.393	0.25
Opportunity for face-to-face time with instructor	During instruction	21	0.298***	0.074	0.154	0.442	
	Before or after instruction	11	0.050	0.118	-0.181	0.281	3.70
	Absent/Not reported	18	0.150	0.091	-0.028	0.327	
Opportunity for face-to-face time with peers	During instruction	21	0.300***	0.072	0.159	0.442	
	Before or after instruction	12	0.001	0.111	-0.216	0.218	5.20
	Absent/Not reported	17	0.184*	0.093	0.001	0.367	
Opportunity to practice	Present	41	0.212***	0.056	0.102	0.322	
	Absent/Not reported	9	0.159	0.124	-0.084	0.402	0.15
Feedback provided	Present	23	0.204**	0.078	0.051	0.356	
	Absent/Not reported	27	0.203**	0.070	0.066	0.339	0.00

[a] The moderator analysis for this variable excluded studies that did not report information for this feature.
*$p < .05$. **$p < .01$. ***$p < .001$.

The test of the practice variable most central to this study—whether a blended online condition including face-to-face elements is associated with greater advantages over classroom instruction than is purely online learning—was discussed earlier. As noted there, the effect size for blended approaches contrasted against face-to-face instruction is larger than that for purely online approaches contrasted against face-to-face instruction.

The other practice variables included in the conceptual framework were tested in a similar fashion. Pedagogical approach was found to moderate significantly the size of the online learning effect ($Q = 6.19$, $p < .05$). The mean effect size for collaborative instruction (+0.25), as well as that for expository instruction (+0.39), was significantly positive, whereas the mean effect size for independent, active online learning (+0.05) was not. Among the other 11 practices, none attained statistical significance. The amount of time that students in the treatment condition spent on task compared with students in the face-to-face condition did approach statistical significance as a moderator of effectiveness ($Q = 3.62$, $p = .06$). The mean effect size for studies with more time spent on task by online learners than learners in the control condition was +0.45, compared with +0.18 for studies in which the learners in the face-to-face condition spent as much or more time on task.

Failure to find significance of most of the coded practices may be a function of limited power after removing studies that did not report what was done with respect to the practice. For example, the synchronicity of computer-mediated communication with the instructor available to online students was documented for only 24 of the 50 contrasts in the meta-analysis. For those 24 contrasts, the size of the effect did not vary significantly between studies in which communication was purely asynchronous and those in which both synchronous and asynchronous communication were available ($Q = 1.20$, $p > .05$). Other practice variables, such as treatment duration, were coded at a relatively coarse level (less than one month vs. a month or more), and future research may uncover duration-related influences on effectiveness by examining more extreme values (for example, a year or more of online learning compared with brief episodes).

CONDITION VARIABLES

The strategy to investigate whether study effect sizes varied with publication year, which was taken as a proxy for the sophistication of available technology, involved splitting the study sample into two subsets by contrasting studies published between 1996 and 2003 against those published

in 2004 through July 2008. Publication period did not moderate the effectiveness of online learning significantly.

To investigate whether online learning is more advantageous for some types of learners than for others, the studies were divided into three subsets of learner type: K–12 students, undergraduate students (the largest single group), and other types of learners (graduate students or individuals receiving job-related training). As noted previously, the studies covered a wide range of subjects, but medicine and health care were the most common. Accordingly, these studies were contrasted against studies in other fields. Neither learner type nor subject area emerged as a statistically significant moderator of the effectiveness of online learning. In summary, for the range of student types for which controlled studies are available, online learning appeared more effective than traditional face-to-face instruction in both older and newer studies, with both younger and older learners, and in both medical and other subject areas. Table 6 provides the results of the analysis of these variables.

Table 6. Tests of Conditions as Moderator Variables

Variable	Contrast	Number Studies	Weighted Effect Size	Standard Error	Lower Limit	Upper Limit	Q-Statistic
Year Published	1997–2003	13	0.195	0.105	-0.010	0.400	0.00
	2004 or after	37	0.203***	0.058	0.088	0.317	
Learner Type	K–12 students	7	0.1664	0.118	-0.065	0.397	3.25
	Undergraduate	21	0.309***	0.083	0.147	0.471	
	Graduate student/Other	21	0.100	0.084	-0.064	0.264	
Subject Matter	Medical/ Health care	16	0.205*	0.090	0.028	0.382	0.00
	Other	34	0.199**	0.062	0.0770	0.320	

*$p < .05$. **$p < .01$. ***$p < .001$.

METHODS VARIABLES

The advantage of meta-analysis is its ability to uncover generalizable effects by looking across a range of studies that have operationalized the construct under study in different ways, studied it in different contexts, and used different methods and outcome measures. However, the inclusion of poorly designed and small-sample studies in a meta-analysis corpus raises concern because doing so may give undue weight to spurious effects. Study methods variables were examined as potential moderators to explore this issue. The results are shown in Table 7.

The influence of study sample size was examined by dividing studies into three subsets, according to the number of learners for which outcome data were collected. Sample size was not found to be a statistically significant moderator of online learning effects. Thus, there is no evidence that the inclusion of small-sample studies in the meta-analysis was responsible for the overall finding of a positive outcome for online learning.

Comparisons of the three designs deemed acceptable for this meta-analysis (random-assignment experiments, quasi-experiments with statistical control, and crossover designs) indicate that study design is not significant as a moderator variable (see Table 7). Moreover, in contrast with early meta-analyses in computer-based instruction and web-based training, in which effect size was inversely related to study design quality (Pearson et al., 2005; Sitzmann et al., 2006), those experiments that used random assignment in the present corpus produced significantly positive effects ($+0.25$, $p < .001$), whereas the quasi-experiments and crossover designs did not (both $p > .05$).

Table 7. Tests of Study Features as Moderator Variables

Variable	Contrast	Number Studies	Weighted Effect Size	Standard Error	Lower Limit	Upper Limit	Q-Statistic
Sample size	Fewer than 35	11	0.203	0.139	-0.069	0.476	0.01
	From 35 to 100	20	0.209*	0.086	0.039	0.378	
	More than 100	19	0.199**	0.072	0.058	0.339	
Type of knowledge tested[a]	Declarative	12	0.180	0.097	-0.010	0.370	0.37
	Procedural/ Procedural and declarative	30	0.239***	0.068	0.106	0.373	
	Strategic knowledge	5	0.281	0.168	-0.047	0.610	
Study design	Random assignment control	32	0.249***	0.065	0.122	0.376	1.50
	Quasi-experimental design with statistical control	13	0.108	0.095	-0.079	0.295	
	Crossover design	5	0.189	0.158	-0.120	0.499	
Unit of assignment to conditions[a]	Individual	32	0.169*	0.066	0.040	0. 298	4.73
	Class section	7	0.475***	0.139	0.202	0.748	
	Course/School	9	0.120	0.103	-0.083	0.323	
Instructor equivalence[a]	Same instructor	20	0.176*	0.078	0.024	0.329	0.73
	Different instructor	19	0.083	0.077	-0.067	0.233	
Equivalence of curriculum/ instruction[a]	Identical/ Almost identical	29	0.130*	0.063	0.007	0.252	6.85**
	Different/ Somewhat different	17	0.402***	0.083	0.239	0.565	

[a] The moderator analysis excluded some studies because they did not report information about this feature.
*$p < .05$. **$p < .01$. ***$p < .001$.

The only study method variable that proved to be a significant moderator of effect size was comparability of the instructional materials and approach for treatment and control students. The analysts coding study features examined the descriptions of the instructional materials and the instructional approach for each study and coded them as "identical," "almost identical," "different," or "somewhat different" across conditions. Adjacent coding categories were combined (creating the two study subsets identical/almost identical and different/somewhat different) to test equivalence of curriculum/instruction as a moderator variable. Equivalence of curriculum/instruction was a significant moderator variable ($Q = 6.85$, $p < .01$). An examination of the study subgroups shows that the average effect for studies in which online learning and face-to-face instruction were described as identical or nearly so was +0.13, $p < .05$, compared with an average effect of +0.40 ($p < .001$) for studies in which curriculum materials and instructional approach varied more substantially across conditions.

Effect sizes did not vary depending on whether or not the same instructor or instructors taught in the face-to-face and online conditions ($Q = 0.73$, $p > .05$) or depending on the type of knowledge tested ($Q = 0.37$, $p > .05$).

The moderator variable analysis for aspects of study method did find some patterns in the data that did not attain statistical significance but that should be retested once the set of available rigorous studies of online learning has expanded. The unit assigned to treatment and control conditions fell just short of significance as a moderator variable ($Q = 4.73$, $p < .10$). Effects tended to be smaller in studies in which whole courses or schools were assigned to online and face-to-face conditions than in those in which course sections or individual students were assigned to conditions.

DISCUSSION AND IMPLICATIONS

The corpus of 50 effect sizes extracted from 45 studies meeting meta-analysis inclusion criteria was sufficient to demonstrate that in recent applications, purely online learning has been equivalent to face-to-face instruction in effectiveness, and blended approaches have been more effective than instruction offered entirely in face-to-face mode.

The test for homogeneity of effects found significant variability in the effect sizes for the different online learning studies, justifying a search for moderator variables that could explain the differences in outcomes. The moderator variable analysis found only three moderators significant at $p < .05$. Effects were larger when a blended rather than a purely online

condition was compared with face-to-face instruction; when the online pedagogy was expository or collaborative rather than independent in nature; and when the curricular materials and instruction varied between the online and face-to-face conditions. This pattern of significant moderator variables is consistent with the interpretation that the advantage of online conditions in these recent studies stems from aspects of the treatment conditions other than the use of the Internet for delivery per se.

Clark (1983) has cautioned against interpreting studies of instruction in different media as demonstrating an effect for a given medium inasmuch as conditions may vary with respect to a whole set of instructor and content variables. That caution applies well to the findings of this meta-analysis, which should not be construed as demonstrating that online learning is superior as a medium. Rather, it is the combination of elements in the treatment conditions, especially the inclusion of different kinds of learning activities, that has proved effective across studies. Studies using blended learning tended also to involve more learning time, additional instructional resources, and course elements that encourage interactions among learners. This confounding leaves open the possibility that one or all of these other practice variables, rather than the blending of online and offline media per se, accounts for the particularly positive outcomes for blended learning in the studies included in the meta-analysis. From a practical standpoint, however, a major reason for using blended learning approaches is to increase the amount of time that students spend engaging with the instructional materials. The meta-analysis findings do not support simply putting an existing course online, but they do support redesigning instruction to incorporate additional learning opportunities online while retaining elements of face-to-face instruction. The positive findings with respect to blended learning approaches documented in the meta-analysis provide justification for the investment in the development of blended courses.

Several practices and conditions associated with differential effectiveness in distance education meta-analyses (e.g., the use of one-way video or audio, computer-mediated communication with instructor) were not found to be significant moderators of effects in this meta-analysis of web-based online learning, nor did tests for the incorporation of instructional elements of computer-based instruction (e.g., online practice opportunities and feedback to learners) find that these variables made a difference. Online learning conditions produced better outcomes than face-to-face learning alone, regardless of whether these instructional practices were used. The implication here is that the field does not yet have a set of instructional design principles sufficiently powerful to yield consistent advantage. Much of the literature on how to implement online or

blended learning (e.g., Bersin, 2004; Martyn, 2003) is based either on interpretations drawn from theories of learning or on common practice rather than on empirical evidence.

The meta-analysis did not find differences in average effect size between studies published before 2004 (which might have used less sophisticated web-based technologies than those available since) and studies published from 2004 on (possibly reflecting the more sophisticated graphics and animations or more complex instructional designs available). However, there were not enough studies in the corpus to test finer-grained categories of online technology (e.g., use of shared graphical whiteboards), nor were differences associated with the nature of the subject matter involved.

Finally, the examination of the influence of study method variables found that effect sizes did not vary significantly with study sample size or with type of design. It is reassuring to note that, on average, online learning produced better student learning outcomes than face-to-face instruction in those studies with random-assignment experimental designs ($p < .001$) and in those studies with the largest sample sizes ($p < .001$).

The relatively small number of studies featuring some of the practices and conditions of interest that also met the basic criteria for inclusion in a meta-analysis limited the power of tests for many of the moderator variables. Some of the contrasts that did not attain significance (e.g., relative time on task and type of knowledge tested) may prove significant when tested in future meta-analyses with a larger corpus of studies.

Meta-analyses are valuable tools for characterizing the evidence base for an educational practice objectively, but they have their limitations. Meta-analyses are always subject to criticism on the grounds that studies of different versions of the phenomenon have been grouped together for quantitative synthesis (Bethel & Bernard, 2010). Some researchers will want to examine only those studies of full-course interventions, only those studies involving K–12 students, only those studies of online mathematics learning, and so on. For some study subsets, results are likely to vary from the overall pattern reported here. However, there are very few controlled empirical studies in most of these subsets, and researchers need to be concerned about basing conclusions on such a narrow base.

Meta-analyses of specific kinds of online learning or for specific learner populations will become desirable, however, as the number of controlled studies in these areas increases. Fortunately, the body of available research for specific kinds of students and learning content and circumstances can be expected to grow rapidly as online learning initiatives continue to expand. Given the growing use of online options with precollege students, and especially for credit recovery programs, there is a particu-

larly urgent need for more well-designed studies of alternative models for younger and less advanced students. Policy makers and practitioners require a better understanding of the kinds of online activities and teacher supports that enable these students to learn effectively in online environments. Future experimental and controlled quasi-experimental designs should report the practice features of both experimental and control conditions to support future meta-analyses of the effectiveness of alternative online learning approaches for specific types of students. Effective practices for learners with different levels of motivation and different senses of efficacy in the subject domain of the online experience need to be studied as well.

Even with this expected expansion of the research base, however, meta-analyses of online learning effectiveness studies will remain limited in several respects. Inevitably, they do not reflect the latest technology innovations. The cycle time for study design, execution, analysis, and publication cannot keep up with the fast-changing world of Internet technology. In the present case, important technology practices of the last five years, notably the use of social networking technology to create online study groups and recommend learning resources, are not reflected in the corpus of published studies included in this meta-analysis.

In addition, meta-analyses of effectiveness studies provide only limited guidance for instructional design and implementation. Moderator variable analyses, such as those reported here, yield reasonable hypotheses as to factors that can influence the effectiveness of online instruction but offer only general guidance to those engaged in developing purely online or blended learning experiences. Feature coding of large numbers of studies to support moderator variable analysis necessarily sacrifices detailed description and context. Meta-analysis is better suited to answering questions about whether to consider implementing online learning or what features to look for in judging online learning products than to guiding the myriad of decisions involved in actually designing and implementing online learning.

We expect more well-designed studies of alternative online learning models to emerge as this kind of instruction becomes increasingly mainstream. But instructional design involves thousands, if not millions, of decisions about the details of structuring a learner's engagement with the material to be learned. Moreover, some studies are finding that design principles that have empirical support when applied to some kinds of learning content prove ineffective with other content (Wylie, Koedinger, & Mitamura, 2009). Under these circumstances, the resources available for online learning research are sure to be outstripped by the sheer number of decisions to be made. Other research approaches, in which online

learning research and development activities are interwoven, are starting to emerge within education (Feng, Heffernan, & Koedinger, 2010). The U.S. Department of Technology's 2010 Education Technology Plan (U.S. Department of Education, 2010), for example, highlights the potential for gaining insights from mining fine-grained learner interaction data collected by online systems. However, these approaches have yet to develop systematic techniques for achieving what meta-analysis of experimental studies does do well—systematically and objectively combining research findings across systems to build a robust knowledge base.

Acknowledgments

Notes

1.   A study by Ellis, Wood, and Thorpe (2004) provides an example of conditions that we did not consider online learning based on this criterion. In this study, three instructional delivery conditions were compared: (1) traditional face-to-face instruction, (2) face-to-face instruction supplemented with printed materials, audio, video, software, and email, and (3) independent study with a CD-ROM, which included "a virtual learning environment"

(p. 359). The second condition includes the use of emails, but email exchanges were not described as a major component of the instruction. The virtual learning environment in the third condition was not offered over the Internet.

2. After completion of this meta-analysis, in response to a suggestion from one of this article's reviewers, we looked at the studies of online learning listed on the NSD (No Significant Difference) website established to document studies that have found no difference in learning outcomes based on the modality of delivery (http://www.nosignificantdifference.org/about.asp). A search with the term *online* generated 36 studies sorted by their conclusion, of which 26 reported no significant difference (10 reported a significant difference, with 9 of these differences favoring the online condition). A quick review of the nature of the NSD studies found that 8 of the listed articles were not empirical studies, 16 lacked control for potential preexisting group differences, 4 lacked an objective measure of student learning, 2 were duplicates of studies listed earlier on the NSD website, 4 were not available online for review, and 2 were outside the time frame for our meta-analysis.

References

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

*Aberson, C. L., Berger, D. E., & Romero, V. L. (2003). Evaluation of an interactive tutorial for teaching hypothesis testing concepts. *Teaching of Psychology, 30*(1), 75–78.

*Al-Jarf, R. S. (2004). The effects of Web-based learning on struggling EFL college writers. *Foreign Language Annals, 37*(1), 49–57.

Allen, I. E., & Seaman, J. (2003). *Sizing the opportunity: The quality and extent of online education in the United States, 2002 and 2003.* Retrieved from http://sloanconsortium.org/publications/survey/sizing_the_opportunity2003

Allen, I. E., & Seaman, J. (2010). *Learning on demand : Online education in the United States, 2009.* Retrieved from http://www.sloanc.org/publications/survey/pdf/learningon demand.pdf

Barab, S. A., Squire, K., & Dueber, B. (2000). Supporting authenticity through participatory learning. *Educational Technology Research and Development, 48*(2), 37–62.

Barab, S. A., & Thomas, M. K. (2001). Online learning: From information dissemination to fostering collaboration. *Journal of Interactive Learning Research, 12*(1), 105–143.

Bates, A. W. (1997). The future of educational technology. *Learning Quarterly, 2,* 7–16.

*Beeckman, D., Schoonhoven, L., Boucque, H., Van Maele, G., & Defloor, T. (2008). Pressure ulcers: E-learning to improve classification by nurses and nursing students. *Journal of Clinical Nursing, 17*(13), 1697–1707.

*Bello, G., Pennisi, M. A., Maviglia, R., Maggiore, S. M., Bocci, M. G., Montini, L., & Antonelli, M. (2005). Online vs. live methods for teaching difficult airway management to anesthesiology residents. *Intensive Care Medicine, 31*(4), 547–552.

*Benjamin, S. E., Tate, D. F., Bangdiwala, S. I., Neelon, B. H., Ammerman, A. S., Dodds, J. M., & Ward, D. S. (2008). Preparing child care health consultants to address childhood overweight: A randomized controlled trial comparing web to in-person training. *Maternal and Child Health Journal, 12*(5), 662–669.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P. A., . . . Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379–439.

Bersin, J. (2004). *The blended learning book: best practices, proven methodologies, and lessons learned.* San Francisco, CA: Wiley.

Bethel, E. C., & Bernard, R. M. (2010). Developments and trends in synthesizing diverse forms of evidence: Beyond comparisons between distance education and classroom instruction. *Distance Education, 31*(3), 231–256.

*Beyea, J. A., Wong, E., Bromwich, M., Weston, W. W., & Fung, K. (2008). Evaluation of a particle repositioning maneuver web-based teaching module. *The Laryngoscope, 118*(1), 175–180.

Bhattacharya, M. (1999). A study of asynchronous and synchronous discussion on cognitive maps in a distributed learning environment. In *Proceedings of WebNet World Conference on the WWW and Internet 1999* (pp. 100–105). Chesapeake, VA: AACE.

Biostat Solutions. (2006). *Comprehensive Meta-Analysis* (Version 2.2.027). Mt. Airy, MD: Biostat Solutions.

Bonk, C. J., & Graham, C. R. (Eds.). (2005). *Handbook of blended learning: Global perspectives, local designs.* San Francisco, CA: Pfeiffer.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Chichester, England: Wiley.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

*Caldwell, E. R. (2006). *A comparative study of three instructional modalities in a computer programming course: Traditional instruction, web-based instruction, and online instruction* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. AAT 3227694)

Cavanaugh, C. (2001). The effectiveness of interactive distance education technologies in K–12 learning: A meta-analysis. *International Journal of Educational Telecommunications, 7*(1), 73–78.

Cavanaugh, C., Gillan, K. J., Kromrey, J., Hess, M., & Blomeyer, R. (2004). *The effects of distance education on K–12 student outcomes: A meta-analysis.* Retrieved from http://www.ncrel.org/tech/distance/index.html

*Cavus, N., Uzonboylu, H., & Ibrahim, D. (2007). Assessing the success rate of students using a learning management system together with a collaborative tool in web-based teaching of programming languages. *Journal of Educational Computing Research, 36*(3), 301–321.

Childs, J. M. (2001). *Digital skill training research: Preliminary guidelines for distributed learning* (Final report). Retrieved from http://www.stormingmedia.us/24/2471/A247193.htm

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disrupting innovation will change the way the world learns.* New York, NY: McGraw-Hill.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445–449.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

*Davis, J. D., Odell, M., Abbitt, J., & Amos, D. (1999, March). *Developing online courses: A comparison of web-based instruction with traditional instruction.* Paper presented at the Society for Information Technology & Teacher Education International Conference, Chesapeake, Va. Retrieved from http://www.editlib.org/INDEX.CFM?fuseaction=Reader.ViewAbstract&paper_id=7520

*Day, T. M., Raven, M. R., & Newman, M. E. (1998). The effects of World Wide Web instruction and traditional instruction and learning styles on achievement and changes in student attitudes in a technical writing in agricommunication course. *Journal of Agricultural Education, 39*(4), 65–75.

*DeBord, K. A., Aruguete, M. S., & Muhlig, J. (2004). Are computer-assisted teaching methods effective? *Teaching of Psychology, 31*(1), 65–68.

Dede, C. (2000). *The role of emerging technologies for knowledge mobilization, dissemination, and use in education.* Paper commissioned by the Office of Educational Research and Improvement, U.S. Department of Education.

Dede, C. (Ed.). (2006). *Online professional development for teachers: Emerging models and methods.* Cambridge, MA: Harvard Education Publishing Group.

Dynarski, M., Agodini, R., Heavisude, S., Novak, T., Carey, N., Campuzano, L., . . . Sussex, W. (2007). *Effectiveness of reading and mathematics software products: Findings from the first student cohort.* Report to Congress. NCEE 2007-4006. Washington, DC: U.S. Department of Education.

*El-Deghaidy, H., & Nouby, A. (2008). Effectiveness of a blended e-learning cooperative approach in an Egyptian teacher education programme. *Computers & Education, 51*(3), 988–1006.

Ellis, R. C. T., Wood, G. D., & Thorpe, T. (2004). Technology-based learning and the project manager. *Engineering, Construction and Architectural Management, 11*(5), 358–365.

*Englert, C. S., Zhao, Y., Dunsmore, K., Collings, N. Y., & Wolbers, K. (2007). Scaffolding the writing of students with disabilities through procedural facilitation: Using an Internet-based technology to improve performance. *Learning Disability Quarterly, 30*(1), 9–29.

Feng, M., Heffernan, N.T., & Koedinger, K.R. (2010). Using data mining findings to aid searching for better cognitive models. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the International Conference on Intelligent Tutoring Systems* (pp. 368–370). Heidelberg, Berlin, Germany: Springer.

Florida TaxWatch. (2007). *Final report: A comprehensive assessment of Florida virtual school.* Retrieved from http://www.floridataxwatch.org/resources/pdf/110507FinalReport FLVS.pdf

*Frederickson, N., Reed, P., & Clifford, V. (2005). Evaluating Web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education, 50*(4), 645–664.

Galvis, A. H., McIntyre, C., & Hsi, S. (2006). *Framework for the design and delivery of effective global blended learning experiences.* Report prepared for the World Bank Group, Human Resources Leadership and Organizational Effectiveness Unit. Unpublished manuscript.

*Gilliver, R. S., Randall, B., & Pok, Y. M. (1998). Learning in cyberspace: Shaping the future. *Journal of Computer Assisted Learning, 14*(3), 212–222.

Graham, C. R. (2005). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 3–21). San Francisco, CA: Pfeiffer.

*Hairston, N. R. (2007). *Employees' attitudes toward e-learning: Implications for policy in industry environments* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. AAT 3257874)

*Harris, J. M., Elliott, T. E., Davis, B. E., Chabal, C., Fulginiti, J. V., & Fine, P. G. (2008). Educating generalist physicians about chronic pain: Live experts and online education can provide durable benefits. *Pain Medicine, 9*(5), 555–563.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hermann, F., Rummel, N., & Spada, H. (2001). Solving the case together: The challenge of net-based interdisciplinary collaboration. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning* (pp. 293–300). Proceedings of the European conference on computer-supported collaborative learning. Maastricht, The Netherlands: McLuhan Institute.

Horn, M. B., & Staker, H. (2011). *The rise of K-12 blended learning*. Innosight Institute. Retrieved from http://www.innosightinstitute.org/mediaroom/publications/education-publications

*Hugenholtz, N. I. R., de Croon, E. M., Smits, P. B., van Dijk, F. J. H., & Nieuwenhuijsen, K. (2008). Effectiveness of e-learning in continuing medical education for occupational physicians. *Occupational Medicine, 58*(5), 370–372.

*Jang, K. S., Hwang, S. Y., Park, S. J., Kim, Y. M., & Kim, J. (2005). Effects of a Web-based teaching method on undergraduate nursing students' learning of electrocardiography. *Journal of Nursing Education, 44*(1), 35–39.

Jonassen, D. H., Lee, C. B., Yang, C.C., & Laffey, J. (2005). The collaboration principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 247–270). New York, NY: Cambridge University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage.

*Long, M., & Jennings, H. (2005). *"Does it work?":The impact of technology and professional development on student achievement*. Calverton, MD: Macro International.

*Lowry, A. E. (2007). *Effects of online versus face-to-face professional development with a team-based learning community approach on teachers' application of a new instructional practice* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. AAT 3262466)

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

*Maki, W. S., & Maki, R. H. (2002). Multimedia comprehension skill predicts differential outcomes of Web-based and lecture courses. *Journal of Experimental Psychology: Applied, 8*(2), 85–98.

Martyn, M. (2003). The hybrid online model: Good practice. *Educause Quarterly*. Retrieved from http://www.educause.edu/ir/library/pdf/EQM0313.pdf

*Mentzer, G. A., Cryan, J., & Teclehaimanot, B. (2007). A comparison of face-to-face and Web-based classrooms. *Journal of Technology and Teacher Education, 15*(2), 233–246.

*Midmer, D., Kahan, M., & Marlow, B. (2006). Effects of a distance learning program on physicians' opioid- and benzodiazepine-prescribing skills. *Journal of Continuing Education in the Health Professions, 26*(4), 294–301.

National Survey of Student Engagement. (2008). *Promoting engagement for all students: the imperative to look within*. Bloomington: Indiana University, Center for Postsecondary Research. Retrieved from http://www.nsse.iub.edu

*Nguyen, H. Q., Donesky-Cuenco, D., Wolpin, S., Reinke, L. F., Benditt, J. O., Paul, S. M., & Carrieri-Kohlman, V. (2008). Randomized controlled trial of an Internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: Pilot study. *Journal of Medical Internet Research*. Retrieved from http://www.jmir.org/2008/2/e9/

*Ocker, R. J., & Yaverbaum, G. J. (1999). Asynchronous computer-mediated communication versus face-to-face collaboration: Results on student learning, quality and satisfaction. *Group Decision and Negotiation, 8*(5), 427–440.

*O'Dwyer, L. M., Carey, R., & Kleiman, G. (2007). A study of the effectiveness of the Louisiana Algebra I online course. *Journal of Research on Technology in Education, 39*(3), 289–306.

*Padalino, Y., & Peres, H. H. C. (2007). E-learning: A comparative study for knowledge apprehension among nurses. *Revista Latino-Americana de Enfermagem, 15*, 397–403.

Paradise, A. (2008). *2007 State of the industry report*. Alexandria, VA: American Society of Training and Development.

Parsad, B., & Lewis, L. (2008). *Distance education at degree-granting postsecondary institutions: 2006-07*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Pearson, P. D., Ferdig, R. E., Blomeyer, R. L., Jr., & Moran, J. (2005). *The effects of technology on reading performance in the middle school grades: A meta-analysis with recommendations for policy*. Naperville, IL: Learning Point Associates.

*Peterson, C. L., & Bond, N. (2004). Online compared to face-to-face teacher preparation for learning standards-based planning skills. *Journal of Research on Technology in Education, 36*(4), 345–361.

Picciano, A. G., & Seaman, J. (2007). *K–12 online learning: A survey of U.S. school district administrators*. Retrieved from http://www.sloanc.org/publications/survey/K-12_06.asp

Picciano, A. G., & Seaman, J. (2008). *Staying the course: Online education in the United States*. Retrieved from http://www.sloanc.org/publications/survey/pdf/staying_the_course.pdf

Riel, M., & Polin, L. (2004). Online communities: Common ground and critical differences in designing technical environments. In S. A. Barab, R. Kling, & J. H. Gray (Ed.), *Designing for virtual communities in the service of learning* (pp. 16–50). Cambridge, England: Cambridge University Press.

*Rockman et al. (2007). *ED PACE final report*. Submitted to the West Virginia Department of Education. Retrieved from http://www.rockman.com/projects/146.ies.edpace/final-report

Rooney, J. E. (2003). Blending learning opportunities to enhance educational programming and meetings. *Association Management, 55*(5), 26–32.

Rudestam, K. E., & Schoenholtz-Read, J. (2010). The flourishing of adult online education: An overview. In K. E. Rudestam, & J. Schoenholtz-Read (Ed.), *Handbook of online learning* (pp. 1–18). Los Angeles, CA: Sage.

*Schilling, K., Wiecha, J., Polineni, D., & Khalil, S. (2006). An interactive web-based curriculum on evidence-based medicine: Design and effectiveness. *Family Medicine, 38*(2), 126–132.

*Schmeeckle, J. M. (2003). Online training: An evaluation of the effectiveness and efficiency of training law enforcement personnel over the Internet. *Technology, 12*(3), 205–260.

*Schoenfeld-Tacher, R., McConnell, S., & Graham, M. (2001). Do no harm: A comparison of the effects of online vs. traditional delivery media on a science course. *Journal of Science Education and Technology, 10*(3), 257–265.

Schwen, T. M., & Hara, N. (2004). Community of practice: A metaphor for online design. In S. A. Barab, R. Kling, & J. H. Gray (Eds.), *Designing for virtual communities in the service of learning* (pp. 154–178). Cambridge, England: Cambridge University Press.

Shotsberger, P. G. (1999). Forms of synchronous dialogue resulting from web-based professional development. In J. Price et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 1999* (pp. 1777–1782). Chesapeake, VA: AACE.

*Sexton, J. S., Raven, M. R., & Newman, M. E. (2002). A comparison of traditional and World Wide Web methodologies, computer anxiety, and higher order thinking skills in the inservice training of Mississippi 4-H extension agents. *Journal of Agricultural Education, 43*(3), 25–36.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology, 59*, 623–664.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752–760.

Smith, M. S. (2009). Opening education. *Science, 323*(5910), 89–93.

Smith, R., Clark, T., & Blomeyer, R. (2005) *A synthesis of new research on K-12 online learning.* Naperville, IL: Learning Point Associates.

*Spires, H. A., Mason, C., Crissman, C., & Jackson, A. (2001). Exploring the academic self within an electronic mail environment. *Research and Teaching in Developmental Education, 17*(2), 5–14.

*Suter, W. N., & Perry, M. K. (1997). *Evaluation by electronic mail.* Paper presented at the annual meeting of the Mid-South Educational Research Association, Memphis, TN.

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of research. *Review of Educational Research, 76,* 93–135.

Taylor, J. C. (2001). *Fifth generation distance education.* Paper presented at the 20th ICDE World Conference, Düsseldorf, Germany. Retrieved from http://www.usq.edu.au/users/taylorj/conferences.htm

*Turner, M. K., Simon, S. R., Facemyer, K. C., Newhall, L. M., & Veach, T. L. (2006). Web-based learning versus standardized patients for teaching clinical diagnosis: A randomized, controlled, crossover trial. *Teaching and Learning in Medicine, 18*(3), 208–214.

*Urban, C. Q. (2006). *The effects of using computer-based distance education for supplemental instruction compared to traditional tutorial sessions to enhance learning for students at-risk for academic difficulties* (Doctoral dissertation). George Mason University, Fairfax, VA.

U.S. Department of Education. (2010). *Transforming American education: Learning powered by technology.* National Education Technology Plan 2010. Washington, DC: Author.

Veerman, A., & Veldhuis-Diermanse, E. (2001). Collaborative learning through computer-mediated communication in academic education. In P. Dillenbourg, A. Eurelings, & K. Hakkarainen (Eds.), *European perspectives on computer-supported collaborative learning.* Proceedings of the First European Conference on CSCL. Maastricht, The Netherlands: McLuhan Institute, University of Maastricht.

*Vandeweerd, J.-M. E. F., Davies, J. C., Pichbeck, G. L., & Cotton, J. C. (2007). Teaching veterinary radiography by e-learning versus structured tutorial: A randomized, single-blinded controlled trial. *Journal of Veterinary Medical Education, 34*(2), 160–167. Vrasidas, C., & Glass, G. V. (2004). Teacher professional development: Issues and trends. In C. Vrasidas & G. V. Glass (Eds.), *Online professional development for teachers* (pp. 1–12). Greenwich, CT: Information Age.

*Wallace, P. E., & Clariana, R. B. (2000). Achievement predictors for a computer-applications module delivered online. *Journal of Information Systems Education, 11*(1/2), 13–18.

*Wang, L. (2008). Developing and evaluating an interactive multimedia instructional tool: Learning outcomes and user experiences of optometry students. *Journal of Educational Multimedia and Hypermedia, 17*(1), 43–57.

Watson, J. F. (2008). *Blended learning: The convergence of online learning and face-to-face education.* Retrieved from http://www.inacol.org/resources/promisingpractices/NACOL_PP-BlendedLearning-lr.pdf

Watson, J. F., Gemin, B., Ryan, J., & Wicks, M. (2009). *Keeping pace with K–12 online learning: A review of state-level policy and practice.* Retrieved from http://www.kpk12.com/downloads/KeepingPace09-fullreport.pdf

WestEd with Edvance Research. (2008). *Evaluating online learning: Challenges and strategies for success.* Retrieved from http://evalonline.ed.gov/

What Works Clearinghouse. (2007). *Technical details of WWC-conducted computations.* Washington, DC: Author.

Whitehouse, P. L., Breit, L. A., McCloskey, E. M., Ketelhut, D. J., & Dede, C. (2006). An overview of current findings from empirical research on online teacher professional

development. In C. Dede (Ed.), *Online professional development for teachers: Emerging models and methods* (pp. 13–29). Cambridge, MA: Harvard University Press.

Wise, B., & Rothman, R. (2010, June). *The online learning imperative: A solution to three looking crises in education.* Washington, DC: Alliance for Excellent Education.

Wisher, R. A., & Olson, T. M. (2003). *The effectiveness of web-based training.* Alexandria, VA: U.S. Army Research Institute.

Wylie, R., Koedinger, K. R., & Mitamura, T. (2009, July–August). Is self-explanation always better? The effects of adding self-explanation prompts to an English grammar tutor. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society.* Amsterdam, The Netherlands.

Young, J. (2002). "*Hybrid" teaching seeks to end the divide between traditional and online instruction: By blending approaches, colleges hope to save money and meet students' needs.* Retrieved from http://chronicle.com/free/v48/i28/28a03301.htm

*Zacharia, Z. C. (2007). Comparing and combining real and virtual experimentation: An effort to enhance students' conceptual understanding of electric circuits. *Journal of Computer Assisted Learning, 23*(2), 120–132.

*Zhang, D. (2005). Interactive multimedia-based e-learning: A study of effectiveness. *American Journal of Distance Education, 19*(3), 149–162.

*Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F., Jr. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information and Management, 43*(1), 15–27.

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record, 107*(8), 1836–1884.

Zirkle, C. (2003). Distance education and career and technical education: A review of the research literature. *Journal of Vocational Education Research, 28*(2), 161–181.

BARBARA MEANS directs the Center for Technology in Learning at SRI International. Her research focuses on the interplay between technology, education systems, and student learning. Currently, she is directing the evaluation of 48 Next Generation Learning Challenges grants that are using a range of technology-based approaches designed to improve students' learning experiences, and college readiness and likelihood of completion.

YUKIE TOYAMA is an education researcher at SRI International's Center for Technology in Learning and a doctoral student in quantitative methods and evaluation at the University of California, Berkeley Graduate School of Education. Her research interests include educational measurement and evaluation designs. Her recent work includes evaluation and research on innovative teaching and learning practices supported by technology both in the United States and abroad.

ROBERT MURPHY is a senior research social scientist at SRI International's Center for Technology in Learning. His research focuses

on design and implementation of large-scale experimental and quasi-experimental evaluations of widely adopted educational programs and technologies. Currently he is coleading a quasi-experimental study of virtual schooling in high school and leading two other studies on the use of blended learning models in K–12 school settings.

MARIANNE BAKIA is a senior social science researcher at SRI International's Center for Technology in Learning. She leads large-scale research and evaluation projects related to the use of educational technologies in public schools. Her research interests include the economic analysis of educational technologies and the support of at-risk students in online learning. Her recent work examines variations in student outcomes in online courses.