Journal of Teacher Education
64(5) 426–438

DOI: 10.1177/0022487113494413
jte.sagepub.com
SAGE

Comparing the Impact of Online and Face-to-Face Professional Development in the Context of Curriculum Implementation

Barry Fishman[1], Spyros Konstantopoulos[2], Beth W. Kubitskey[3], Richard Vath[1], Gina Park[1], Heather Johnson[4], and Daniel C. Edelson[5]

Abstract

This study employed a randomized experiment to examine differences in teacher and student learning from professional development (PD) in two modalities: online and face-to-face. The study explores whether there are differences in teacher knowledge and beliefs, teacher classroom practice, and student learning outcomes related to PD modality. Comparison of classroom practice and student learning outcomes, normally difficult to establish in PD research, is facilitated by the use of a common set of curriculum materials as the content for PD and subsequent teaching. Findings indicate that teachers and students exhibited significant gains in both conditions, and that there was no significant difference between conditions. We discuss implications for the delivery of teacher professional learning.

Keywords

professional development, technology, teacher knowledge, science education

Research on educational outcomes suggests that teachers are a key “variable” in student success (e.g., Nye, Konstantopoulos, & Hedges, 2004), and recent research focuses on precisely how much difference effective teachers have in terms of year-over-year student learning outcomes (e.g., Chetty, Friedman, & Rockoff, 2012). How teachers teach matters tremendously for students’ learning, and given high-quality professional learning opportunities, all teachers have the capacity to improve students’ learning outcomes. Questions about what types of professional development (PD) are most effective and in what contexts are thus critical policy issues. The recent growth in online PD makes the range of choices available for PD more complex, increasing the need for empirical research to support decision making about how to invest in PD. The study reported here explores the overarching research question:

> *Research Question*: How does online PD compare with face-to-face PD in terms of effects on teachers and students when the PD content is held constant?

We examined this question through three subquestions:

> *Research Subquestion 1 (RQ1)*: Are there differences in teachers’ learning in terms of changes in beliefs and knowledge as a function of different PD modalities?
>
> *Research Subquestion 2 (RQ2)*: Are there differences in teachers’ classroom practice?
>
> *Research Subquestion 3 (RQ3)*: Are there differences in student learning outcomes as a function of PD modalities?

In short, we ask which PD modality is most effective when comparing between online and face-to-face, all other things being equal.

The question of what kind(s) of PD are most effective is, of course, unanswerable in the abstract. We also must ask, “Most effective for *what purpose*?” In this study, we focus on teacher learning to support new curriculum adoption. Many school improvement programs are based on interventions designed to support student learning, and curriculum is a core component of teaching and learning contexts (Cohen,

[1]University of Michigan, Ann Arbor, MI, USA
[2]Michigan State University, East Lansing, MI, USA
[3]Eastern Michigan University, Ypsilanti, MI, USA
[4]Vanderbilt University, Nashville, TN, USA
[5]The National Geographic Society, Washington, DC, USA

Corresponding Author:
Barry Fishman, 610 E. University, Room 4121, University of Michigan, Ann Arbor, MI 48109, USA.
Email: fishman@umich.edu

Raudenbush, & Ball, 2003). Teachers play a key moderating role in how materials are interpreted and presented to students (Barab & Luehmann, 2003), and PD is key to helping teachers build capacity to use curriculum well (Desimone, Porter, Garet, Yoon, & Birman, 2002; Fishman, Marx, Best, & Tal, 2003).

Online PD, which we define as teacher learning experiences delivered partially or completely over the Internet, can potentially provide high-quality teacher learning experiences. Recent U.S. policy documents, such as the 2010 National Educational Technology Plan, advocate strongly for online teacher learning (U.S. Department of Education, 2010). But what do we really understand about the opportunities and challenges presented by online PD? Does online PD work as well as face-to-face PD? This study addresses these questions directly, through an experimental comparison of teachers learning how to implement new secondary science curriculum through either face-to-face or online PD.

We found no difference in outcomes as a function of PD modality. In online and face-to-face PD conditions, teachers reported increased confidence with new curriculum materials, enacted those materials consistently with curriculum designers' intent, and their students learned from curriculum successfully and in equal amounts. Below, we describe this study and its findings in detail, and consider its potential meaning for those who would select, design, and/or deliver learning opportunities for teachers, particularly in the context of curriculum reform.

Literature Review and Conceptual Framework

Our study contributes to two important threads in the literature on teacher PD: (a) An increasing focus on changes in classroom practice and student learning as outcomes of PD, as opposed to focusing only on changes in teachers' knowledge and beliefs, and (b) a trend toward offering teacher PD online to increase availability and scalability.

The Evolution of Research on Teacher Learning from PD

Scholarship on teachers' learning has evolved from studies of attitude and belief to the increasing use of research designs measuring impact on teacher and student learning (Desimone, 2009; Scher & O'Reilly, 2009). Research on PD effectiveness has transitioned from an over-reliance on teachers' self-reported attitudes about PD (Frechtling, Sharp, Carey, & Vaden-Kiernan, 1995) to a greater examination of objective outcomes, utilizing classroom practice and student learning data (Loucks-Horsley & Matsumoto, 1999). Evidence about the differential effects of PD characteristics on teacher practice has emphasized that the most effective PD employs active learning techniques, is coherent with other school initiatives, and is of extended duration (Garet, Porter, Desimone, Birman, & Yoon, 2001).

A persistent challenge for research on learning from PD is the *direct* measurement of impact on practice and student learning outcomes, as opposed to teacher self-report of change. Scholars are increasingly calling for studies that directly document teacher learning, classroom practice, student learning, and the links among them (Desimone, 2009; Fishman et al., 2003). Scher and O'Reilly (2009) conducted a meta-analytic review of studies since 1990 that employed a quasi-experimental or experimental design linking math and science PD to either changes in teacher attitudes, teacher practice, student attitudes, or student learning. Their review identified only 18 studies meeting their criteria, and only a single one of those was a randomized experiment. But the number of studies that focus on all three elements has been increasing. For example, Garet and colleagues (2008) conducted an experiment on the impact of PD coaching interventions on the knowledge and practice of second-grade teachers, and on student reading outcomes. Santagata, Kersting, Givvin, and Stigler (2011) employed a randomized trial to study teachers learning from TIMMS video modules, examining effects on teacher knowledge, classroom implementation, and student mathematics learning.

The Importance of Context for Studying Teacher Learning

PD frequently focuses on generalized instructional practices, such as collaborative learning. But a challenge of such generalized PD is that it is up to each teacher to interpret the PD content and translate it into specific action in their classroom. In addition to being challenging for teachers, this is also challenging for researchers, because there is no common baseline of teaching practice or student learning outcomes that would allow for measurement of effects or comparison across teachers who participated in the PD. But one group of PD providers, curriculum publishers, *do* have an incentive to provide direct support for classroom practice. Publishers want teachers to succeed with their materials, and often provide school districts with PD as a part of the purchase contract for new materials. This provides a valuable entry point for studying teacher learning, as curriculum materials can be a major vehicle for introducing new teaching practices into classrooms (Ball & Cohen, 1996; Davis & Krajcik, 2005). When new and challenging content or ways of organizing instruction are being introduced, such as with inquiry-oriented science, curriculum can itself serve as a form of PD for teachers by providing concrete examples of instructional practice that link learning goals, assessments, and standards (Krajcik, McNeill, & Reiser, 2008).

Studying teacher learning in the context of curriculum adoption represents an opportunity to conduct research that simultaneously evaluates teacher learning, classroom practice, and student learning. The study of what teachers learn from PD in terms of their content knowledge (CK) and beliefs is bounded by the knowledge required to teach the curriculum. Changes in teachers' classroom instructional practice as a function of PD can be studied directly, as the curriculum provides activities that are likely to be enacted, if not identically, at least in ways that resemble each other in different classrooms. The direct comparison of student learning outcomes is enabled by the shared content of the curriculum materials. With these three elements constrained by the curriculum, researchers can vary PD features and treat teacher knowledge, instructional practice, and student learning as outcomes or dependent variables. In the study described in this paper, we examined these three types of outcomes as a function of systematic variation of the modality by which PD was delivered to teachers: in either online or face-to-face workshops.

Opportunities and Issues Presented by Online PD

The rapid growth of the Internet in all aspects of modern life and work makes online PD an attractive choice (O'Dwyer, Carey, & Kleiman, 2007). As Dede, Ketelhut, Whitehouse, and McCloskey (2009) argue, online PD has many potential advantages: It can accommodate teachers' busy schedules, it can draw on powerful resources that are not available locally, and it holds the promise of creating a "path toward providing real-time, work-embedded support" (p. 9) for teachers' ongoing learning. Online PD does not require teachers to assemble in the same location, making it available to geographically isolated teachers. However, there is concern over what might be lost in online PD. Is there a loss in terms of building trust and local collegiality, or providing teachers with hands-on experiences? Teachers can be valuable resources for each others' learning in PD, sharing experiences and guidance on how to enact challenging curriculum or activities in the classroom, and building collegiality is important in facilitating sharing (Kubitskey, Fishman, & Marx, 2002). Do teachers' prior beliefs about web-based learning (Kao & Tsai, 2009) lead some to develop biases against online PD? Can online learning provide "active" learning experiences for teachers (Garet et al., 2001)?

Research on teacher learning from online PD has the same challenges and limitations discussed earlier with respect to "traditional" PD (for broad-based reviews of the literature on online PD, see Dede et al., 2009 and Ginsberg, Gray, & Levin, 2004). While there has been progress in research designs related to face-to-face PD (discussed above), research around online PD, for the most part, echoes the observations of Frechtling et al. (1995). In their call for a better-specified research agenda for online teacher learning, Dede and colleagues conclude that, "Evidence of effectiveness is often lacking, anecdotal, or based on participant surveys completed immediately after the professional development experience" (Dede et al., 2009, p. 9). Though some research addresses teacher pedagogical change, none of the research identified in their review tackles the issue of how online PD relates to student learning outcomes. In this way, online PD scholarship mirrors research on online learning in general, where reviews fail to find well-designed studies focused on learning outcomes (Means, Toyama, Murphy, Bakia, & Jones, 2009).

Studies Comparing Face-to-Face and Online PD

There is an argument in the educational technology literature that, all other things being equal, one should not expect different outcomes when the only change is based on media (Clark, 1983). By "all other things," we refer to the PD pedagogical design, PD content, and intended context for enactment of PD content. A counter argument is offered by Kozma (1994), who argued that the debate about media effects needed to be reframed to ask in what ways the affordances of various media might allow us to support learning. For example, perhaps online PD allows teachers to take their time and focus on what is most important to them, whereas face-to-face PD might better support the exchange of practical experience among teachers. Studies of teachers learning from online PD that employ experimental design with randomization and control groups have started to address the linkages between teachers' learning, practice, and student learning outcomes. We report on the designs and findings of three such studies to set the stage for discussing our own study.

In a study of coaching PD for Head Start teachers, Powell, Diamond, Burchinal, and Koehler (2010) conducted a randomized experiment investigating differences between how technologically mediated remote versus in-person coaching led to differential practice and student literacy outcomes. They found learning and practice improvements in both conditions, with students doing better on some measures in each, but no clear pattern favoring one modality over the other. In the end, the investigators concluded that technology-mediated PD is "a promising alternative to the common coaching practice of in-person visits," and that "Coaching supports to teachers in the remote condition . . . approximated the supports provided to teachers in the on-site condition" (Powell et al., 2010, p. 310).

Masters, Magidin deKramer, O'Dwyer, Dash, and Russell (2010) studied the impact of online PD on the knowledge and instructional practices of fourth-grade English language arts (ELA) teachers in a 7-week PD program. A control group received no PD, but was not prohibited from pursuing normal learning activities. Effects on knowledge were measured using pre- and posttests, with self-report of practice. Findings showed significant effects of the online PD in knowledge growth and practice from pretest to posttest, when compared with the control group. However, because the control group received no organized PD, this was really an evaluation study of the online PD. This study was part of a larger series of four randomized trials with ELA and math teachers in fourth through eighth grades (O'Dwyer et al., 2010), all using a design similar to Masters et al. (2010), but adding student learning measures. Findings from teachers in online conditions were similar with respect to knowledge and self-reported practice. The effects on student learning were not as strong, with significant effects being found on some but not all measures across the different content areas and grade levels.

Fisher, Schumaker, Culbertson, and Deshler (2010) used a randomized controlled trial design to examine teacher learning to use a concept mapping technique to support student learning. Teachers enrolled in a graduate-level course were randomly assigned to receive a virtual or a face-to-face workshop. The virtual workshop consisted of stand-alone training materials on a CD-ROM, allowing teachers to view a lesson plan and interact with virtual coaches and students. The face-to-face workshop design mirrored the content of the virtual workshop. The PD was of relatively short duration, conducted in two sessions. Following the workshops, teachers were tested on their knowledge of the technique and asked about their satisfaction with the workshop. Knowledge growth was significant and positive for both groups, and both groups were satisfied with the PD. No significant differences were detected between the groups on either measure. Eight teachers were observed before and after PD to examine effects on practice and student learning. Implementation of the concept mapping techniques was measured with a checklist, and students' (n = 125) learning was measured with a posttest. Teacher performance and student learning improved in both conditions, with no significant differences detected between groups. The researchers conclude that their study indicates the promise of computerized PD as a valid replacement for face-to-face PD, but acknowledge the small size of the observational study limits its generalizability.

It is encouraging to see the emergence of randomized experimental designs with control groups in the online PD literature. Such studies are critical for examining the opportunities afforded by online PD. Yet no single study can answer all the questions that need to be asked. Studies such as that conducted by Masters et al. (2010) are useful for demonstrating the efficacy of online learning in terms of changing teachers' knowledge and practice, but without a strong comparison condition, they have limited value for guiding policy. Studies such as those conducted by Fisher et al. (2010) and Powell et al. (2010) go further by including well-constructed controls and linking teacher learning to classroom practice and student learning.

The Current Study: Context and PD Design

"Online PD" is not monolithic. It makes little sense to ask questions about whether "it" is more or less effective than any other PD modality. The term describes a vast range of designs, from single-session workshops—that are typically derided in the teacher learning literature—to extended experiences spanning weeks or years. It can describe PD that is completely online or "hybrid" environments that blend face-to-face and online elements (Kleiman, 2004). Thus, when considering questions of comparative effectiveness, it is critical to clearly identify design features of PD opportunities in question. The PD in this study is designed to prepare high school teachers to implement a year-long environmental science curriculum, *Investigations in Environmental Science: A Case-Based Approach to the Study of Environmental Systems* (Edelson et al., 2005), developed with support from the National Science Foundation and licensed to a publisher for commercial distribution.

The curriculum employs a pedagogical approach called Learning-for-Use (Edelson, 2001), designed to develop understanding through cycles of motivation, knowledge construction, and knowledge organization. Learning-for-Use emphasizes iterative construction of understanding through examining problems through multiple lenses over time. The curriculum includes Geographic Information Systems (GIS) and interactive simulations to enable students to investigate science questions. In preparing to implement the curriculum, teachers must extend their knowledge of environmental science, become comfortable with inquiry-based learning in general (Crawford, 2000), and the Learning-for-Use approach in particular, and become familiar with GIS software and how to teach with it. The introduction of this curriculum to a classroom is a robust intervention (Cohen et al., 2003), designed to foster student learning through a highly specified curriculum. But before curriculum *can* influence student learning, teachers must implement it efficaciously (Squire, MaKinster, Barnett, Leuhmann, & Barab, 2003). The PD is intended to increase the probability that teachers' curriculum enactment is consistent with designers' intentions, and leads to desired learning outcomes. The curriculum consists of three units meant to be completed over an academic year.

We developed two PD conditions to accompany teachers' adoption of the curriculum. One condition was a week-long (48 hr) face-to-face workshop spread over 6 days. The other was an online workshop designed to be completed by teachers asynchronously at their own pace. Given the lack of strong empirical data about how to design online PD (c.f., Dede et al., 2009) when we designed our PD in 2005-2006, we relied on a mixture of "best practice" consensus documents (e.g., Loucks-Horsley, Hewson, Love, & Stiles, 1998) and evidence about traditional PD (e.g., Garet et al., 2001). The week-long face-to-face workshop format was chosen because it is common for curricula such as the one in this study. All teachers were brought to a central location for the PD workshop as there was typically a single or just a few environmental science teachers per district (see "Method" section for details on the study participants). The online PD was structured as a series of self-paced "short courses," with a facilitator guiding teachers and answering questions as they worked through materials. The facilitator assisted teachers who got stuck or were confused, encouraged teachers to make steady progress, and gave feedback on teachers' written reflections. Although part of the online materials included a discussion forum, the design of the program neither encouraged nor discouraged teachers to interact with each other as a part of the PD.

To alleviate concerns that teachers might not be successful with the online tools, we convened teachers in the online condition for a 2-day (16 hr) face-to-face orientation session. During this session, teachers were introduced to the online environment and received a general orientation to the curriculum identical to the orientation that the teachers in the face-to-face condition received. Because of this initial face-to-face session, our online condition might be considered "blended" or "hybrid" PD in a strictly technical sense. However, such environments typically involve an interleaving of online and face-to-face environments, or an augmentation of face-to-face with online, whereas our PD environment was completely online after the orientation. Furthermore, the *content* that our data-gathering and analysis focused on was not presented in the orientation session, and was only encountered by teachers in the online materials. For that reason, we do *not* consider the PD experience of online teachers to be blended or hybrid for the purpose of this research. The design of the PD was, we believe, typical for such environments in 2005.

The study was designed to provide the same *opportunities to learn* in each condition, similar to Fisher et al. (2010) and Powell et al. (2010). We designed the PD intervention to ensure the topics were the same, and their presentation was balanced in each condition. All teachers viewed the same computer-based simulations, and had access to the same print-based support materials. However, while working to keep the content constant, we took advantage of the affordances of each condition's medium, without artificially limiting either of the conditions for the sake of comparability. Some ways that the conditions naturally differed included the way(s) teachers interacted with materials, with facilitators, with each other, and the time spent "on task" in PD activities.

The ways teachers interacted with materials varied according to what was most "natural" in a given modality. Where facilitators in the face-to-face condition used PowerPoint to present material, teachers in the online condition would read the same material as text. One does not attend face-to-face PD and expect to sit and read material beside other participants, but online, reading is expected and even preferred in many cases to video presentations. This is because written materials can be skimmed or reviewed easily, whereas video is linear and harder to index or scan.

Teachers working shoulder-to-shoulder for a week tend to develop local community, and questions and discussions related to the core content may emerge that go beyond the original agenda. This is a positive affordance of gathering teachers together at the same time and in the same place, as teachers may recognize their peers' concerns and contributions as more salient than those of an outside "expert," due to the practical classroom experiences shared among teachers (Kubitskey et al., 2002). The online PD was designed to be an individualized experience, where teachers interacted primarily with the materials and with facilitators. We note that our approach to online PD was different from the path taken by some programs, which emphasize building online communities of practice (e.g., Schlager, Fusco, & Schank, 2002). We raise this to emphasize that different design choices may lead to variation in outcomes, a topic we return to in our discussion.

Time spent on task in PD also varied according to condition. In the face-to-face condition, teachers spent 48 hr together over 6 days. As with most face-to-face instruction, the facilitator set the pacing of the session. Time for breaks and meals were part of the 48 hr. While we designed the online PD content to match the face-to-face content, and while our intent was that the online PD should take about the same number of hours as the face-to-face PD, in practice, there is no way to control how quickly or slowly an individual teacher moves through the online materials. We view this as a positive design affordance of online PD, and one that teachers might find appealing. Teachers in the online condition were not limited in terms of when they could interact with the materials, and were free to engage with PD much closer to the time when they enacted related lessons, so that it was fresher for them than it might be for teachers in the face-to-face condition. They may also elect to review materials or spend more time on particular sections of the PD.

Research Questions

Our study focused on the overarching question: "How does online PD compare with face-to-face PD in terms of effects on teachers and students, when the content of the PD is held constant?" We examined this question through three subquestions. The first and second subquestions were related to determining differences between the face-to-face and online PD conditions for important teacher outcomes: (RQ1) "Are there differences in teacher learning in terms of changes in beliefs and knowledge?" To understand this, we examined differences with respect to changes in teacher CK, teacher beliefs about self-efficacy to teach environmental science, and teacher beliefs about teaching science in general. The second question was related to practice: (RQ2) "Are there differences in classroom practice as a function of the PD?" To examine this, we conducted videotaped observations of teacher practice looking for differences in how they taught the curriculum, focusing on key ideas presented in PD. The third question focused on examining differences between the two conditions for student knowledge outcomes from test scores related to the curriculum: (RQ3) "Are there differences in student learning outcomes as a function of the PD?"

Method

Participants

We conducted a cluster randomized experiment where we randomly assigned secondary teachers from across the United States (in schools that adopted the curriculum) to either a face-to-face or online condition; a total of 49 teachers (24 face-to-face and 25 online). The cluster randomized design had two levels: The teacher was the cluster (or second-level unit) and the students who were nested within teachers were the first-level units. As a result, we employed a two-level model (students nested within teachers) to analyze the data (see statistical analysis below). The average number of students per teacher was 23, a total of 1,132 students (522 face-to-face and 610 online). Because our study was designed as a randomized experiment, the results should have high internal validity (i.e., the treatment effect should be unbiased) and causal inferences should be warranted. Our study included 45 schools from urban (n = 6), suburban (n = 22), and rural (n = 17) areas and thus, the results should in principle have higher external validity (i.e., generalizability) than studies with smaller convenience samples.

Measures and Instrumentation

We gathered teacher beliefs and knowledge data using surveys administered online including questions about teachers' background and preparation for teaching science (courses taken and certification status). We measured teachers' environmental science knowledge using a 25-item test of content related to the curriculum. We measured teachers' self-efficacy for teaching environmental science using a modified version of the Science Teacher Efficacy Beliefs Instrument (Riggs & Enochs, 1990) and their feelings of preparedness teaching environmental science. We also asked questions about teachers' general beliefs about the nature of science and of science teaching (scales adapted from Smith, Banilower, McMahon, & Weiss, 2002), including their tendencies toward traditional didactic versus reform-oriented constructivist teaching practice (Becker, 1999). These surveys were administered prior to the initial PD, and again when teachers concluded teaching the curriculum.

To measure teachers' classroom practice in relation to PD, we asked teachers to submit video recordings of predetermined lessons from the curriculum, allowing us to compare practices that were linked to specific content from the PD. Each video served as evidence of teacher enactment of at least one of three strands teachers learned about in PD and that are central to the curriculum: (a) "Making connections" refers to teachers' explicit efforts to support students in making conceptual connections between and across lessons and the larger curriculum context; (b) "Decision making" refers to teachers' explicit efforts to support students in developing skills related to making evidence-based decisions—a central curriculum activity; and (c) "Technology use" refers to teachers' explicit efforts to help students navigate through GIS software to investigate environmental science problems.

We scored each video using rubrics consisting of two main components: quality of curriculum enactment/general teaching and quality of PD enactment. The "quality of curriculum enactment" categories were adapted from Tal, Krajcik, and Blumenfeld's (2006) conception of teachers' role in enacting project-based science curriculum to measure teachers' general pedagogical approaches to enact the curriculum and probe for student understanding, as well as the level of student engagement in the task. In addition, we coded for the ways in which teachers adapted and modified the curriculum. "Quality of PD enactment" categories aim to measure the quality of teachers' enactment of specific strategies teachers were exposed to in PD, which were introduced in comparable ways in both conditions. The categories were developed based on observations and identification of core components discussed in both PD conditions. The scales for each category were developed by watching a subsample of classroom videos to identify variations in teachers' quality of enactment of what they had learned in PD. "Quality of curriculum enactment" categories were generally common across all videos, but "quality of PD enactment" categories were unique to each strand, because they call for different pedagogical approaches and knowledge on the part of the teacher.

Subcategories were scored using binary categorical scales (e.g., whether or not modifications were made to the curriculum, whether or not a specific PD strategy was demonstrated in the lesson) and ordinal scales (e.g., to assess the quality of a specific aspect of enactment) that were summed up or averaged within each category, depending on the type of scale. Videos were reduced to a series of scores representing teachers' enactment of each unique strand addressed in PD. A team of research assistants coded videos, supervised by a "master coder" who led coding scheme development. Prior to coding the videos, reliability of at least 80% was established through training and maintained through random reliability checks throughout the coding process.

Student learning was measured using a 29-item multiple-choice item of environmental science CK developed as a proximal assessment (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002) with high content validity to the curriculum. The assessment was developed and validated as part of a research study conducted in Chicago Public Schools to align inquiry-oriented instruction with high-stakes test outcomes (McGee & Brazdil, 2011).Teachers administered the test prior to the beginning of instruction and then again immediately following instruction. Using a matched pre–posttest design, *student learning* was operationalized as the gain from pre to post (i.e., *posttest raw score* minus *pretest raw score* equals *gain score*).

Teacher Variables

The main independent variable was a binary variable indicating membership in the online condition versus face-to-face condition, which served as the comparison group. Multiple dependent variables represented teacher knowledge, beliefs, and practice outcomes (using instruments discussed above). The first was environmental science CK. Second were teachers' beliefs about their self-efficacy (Riggs & Enochs, 1990), which contained two variables labeled *personal* and *impersonal* efficacy beliefs. The third set of outcomes included measures of teacher beliefs about how science should be taught, containing six variables: (a) attitudes toward teaching science, (b) preparedness to teach science with respect to pedagogical content knowledge (PCK), (c) preparedness to teach science with respect to CK, (d) importance of traditional instructional practices when teaching science, (e) importance of inquiry practices when teaching science, and (f) importance of establishing an investigative culture when teaching science.

When examining teacher practice with respect to PD, we focused on elements of teaching that directly related to the curriculum and were emphasized in the PD activities (using the scheme discussed above). Thus, our fourth set of variables included curriculum enactment, curriculum modification, probing for student understanding, pedagogical strategies, and use of resources (e.g., visual organizers, examples). Fifth, we examined how teachers presented instruction in support of "decision making," an essential component of the curriculum. Our sixth and final group of teacher outcome variables was implementation of curriculum-specified technology such as GIS tools, based on our observations.

Statistical Analysis Approach

For continuous outcomes related to teacher-level data ($n = 49$) such as environmental science CK, beliefs of efficacy in teaching environmental science (personal and impersonal) and beliefs about teaching (attitudes, preparedness, traditional vs. inquiry instructional practices), we used linear regression models. The main independent variable was a binary variable indicating online PD, with face-to-face as the comparison group. In these models, we always controlled for pretest scores as well as for teacher background (e.g., science coursework completed), teaching certificate in environmental science, and years of teaching experience in environmental science. We also controlled for teaching philosophy when we modeled teacher beliefs about teaching environmental science. The teacher outcomes related to making connections, teaching the decision-making process, or using technology were either dichotomous or polytomous (i.e., more than two categories). To examine whether these outcomes were associated with the treatment, we used chi-square tests of independence, which are appropriate for determining the significance of bivariate relationships between categorical variables.

To analyze student data, we used multilevel models (Raudenbush & Bryk, 2002) that take the dependency of students within teachers into account. Specifically, we used a two-level model where students at the first level were nested within teachers at the second level. The main outcome was the CK posttest score. The main independent variable was the binary indicator representing the online condition with face-to-face as the comparison group. The pretest score was used as a covariate at the student level. In addition, gender (e.g., female), race (e.g., Black, Hispanic, Asian, and Other races), GPA and school grade (e.g., 11th and 12th grade) were included as covariates at the student level. The comparison groups for the student-level covariates were male, White, and 9th and 10th grade. The second-level variables included teacher and school-level covariates such as school composition (e.g., percentage of economically disadvantaged students), school locality (e.g., urban and suburban schools with rural schools being the comparison group), teacher experience, teacher certification, teacher science coursework, teacher CK, and beliefs of personal efficacy in teaching environmental science.

Findings

Descriptive Statistics

Table 1 summarizes school, teacher, and student samples descriptive statistics, including means and sample sizes for variables of interest for the total sample and by treatment condition. Across all schools, 20% of students were eligible for free or reduced price lunch. Nearly 50% of schools were in suburban areas, and more than one third of schools were in rural areas. Approximately 80% of teachers were certified in environmental science and had about five years experience in teaching environmental science. The majority of students were males and White. Nearly 30% of students were minority students and approximately 80% of students were in Grades 11 and 12.

Usage Patterns for Teachers in the Online PD Condition

Teachers assigned to the face-to-face condition had 48 hr of summer PD, while teachers in the online condition initially had a 16-hr face-to-face orientation session and, on average, nearly 20 hr of online PD over the next several months. There was considerable range in the amount of time teachers in the online condition spent engaged with online PD, from 3 to 52 hr. Because of the considerable variation, we conducted analyses examining whether time spent online was related to student achievement. The results did not indicate significant association between hours of online activities and gains in

Table 1. Descriptive Statistics of the Sample: Means and Sample Sizes.

Variables	Total sample	n	Online condition	n	Face-to-face condition	n
Schools						
Percent free lunch	0.20	45	0.19	21	0.22	24
Urban	0.13	6	0.14	3	0.13	3
Suburban	0.49	22	0.57	12	0.42	10
Rural	0.38	17	0.29	6	0.46	11
Teachers						
Years of teaching experience: science	10.41	49	9.52	25	11.33	24
Years of teaching experience: environmental science	4.90	49	4.44	25	5.38	24
Certification in environmental science	0.78	49	0.80	25	0.75	24
Students						
Female	0.45	1,089	0.45	596	0.45	493
White	0.71	1,076	0.66	585	0.77	491
Black	0.08	1,076	0.10	585	0.06	491
Hispanic	0.10	1,076	0.10	585	0.10	491
Asian	0.02	1,076	0.03	585	0.01	491
Other race	0.09	1,076	0.11	585	0.05	491
Grade 9	0.16	168	0.25	147	0.04	21
Grade 10	0.05	50	0.05	30	0.04	20
Grade 11	0.33	353	0.28	165	0.39	188
Grade 12	0.46	492	0.41	242	0.52	250
GPA	2.76	1,036	2.82	566	2.69	470

Table 2. Regression Estimates for CK and Personal and Impersonal Efficacy Beliefs.

Variables	CK posttest	Personal efficacy beliefs	Impersonal efficacy beliefs
Online versus face-to-face	0.668 (1.052)	0.183 (0.124)	−0.037 (0.163)
CK pretest	0.176 (0.299)	0.675* (0.187)	0.570* (0.242)
Course work in biological sciences	−0.081 (0.301)	0.061 (0.042)	−0.025 (0.058)
Course work in earth sciences	−0.096 (0.184)	0.007 (0.021)	0.014 (0.029)
Course work in environmental sciences	−0.090 (0.162)	−0.003 (0.018)	−0.005 (0.025)
Certification in environmental sciences	−0.502 (1.410)	−0.078 (0.162)	0.116 (0.222)
Years of teaching experience in environmental science	−0.117 (0.092)	0.003 (0.011)	0.001 (0.015)

Note. Standard errors are in parentheses. CK = content knowledge.
$^{*}p < .05$.

student achievement, our main outcome of interest in the study.

Teacher Outcomes

Teacher knowledge and belief outcomes (RQ1) represent the first of three primary study foci. One teacher outcome of interest was CK in environmental science. The pre to post gain in teacher CK in the online condition was on average 0.88 points and was just below 1/2 of a standard deviation (*SD*). The pre to post gain in teacher CK in the face-to-face condition was on average 0.58 points and was approximately 1/9 of a *SD*. These results suggest teachers in both conditions improved on average with respect to CK. Despite teachers in the online condition having an advantage on average 0.30 points, it was not significant at the .05 level. In the regression models, we regressed the posttest CK scores at the end of the school year on the treatment conditions (online condition was coded 1, and face-to-face condition was coded 0). In the model we also controlled for pretest scores, teacher background (e.g., science coursework completed), teaching certificate in environmental science, and years of teaching experience in environmental science. The results suggested no significant differences in teacher CK between conditions. Thus, it appears that teachers' CK was not affected by condition. The estimates of this analysis are reported in Table 2.

Another outcome of interest was teachers' beliefs about their self-efficacy for teaching environmental science, measured using scales for personal and impersonal beliefs. The pre to post gain in teachers' personal beliefs about personal efficacy in teaching environmental science in the online condition was on average 0.52 points and was nearly one *SD* gain. The

Table 3. Regression Estimates for Teaching Beliefs.

Variables	Attitudes	Investigative culture	Inquiry practices	Prepared CK	Prepared PCK	Traditional practices
Online versus face-to-face	−0.027 (0.100)	0.166 (0.125)	−0.016 (0.109)	0.100 (0.112)	0.031 (0.128)	0.119 (0.098)
CK pretest	0.651* (0.159)	0.313* (0.151)	0.378* (0.121)	0.381* (0.132)	0.443* (0.138)	0.298 (0.156)
Course work in biological sciences	0.005 (0.035)	−0.091* (0.043)	0.084* (0.038)	−0.016 (0.038)	0.012 (0.044)	0.056 (0.034)
Course work in earth sciences	0.003 (0.018)	−0.005 (0.022)	0.031 (0.020)	0.013 (0.020)	0.013 (0.023)	0.016 (0.017)
Course work in environmental sciences	−0.007 (0.015)	−0.014 (0.020)	−0.032 (0.017)	−0.005 (0.017)	−0.012 (0.020)	−0.005 (0.016)
Certification in environmental sciences	0.117 (0.141)	0.051 (0.183)	0.195 (0.156)	0.012 (0.165)	0.183 (0.183)	−0.077 (0.140)
Years of teaching experience in environmental science	0.0005 (0.010)	−0.006 (0.012)	0.015 (0.011)	0.009 (0.011)	−0.002 (0.013)	0.014 (0.010)
Philosophy 1	−0.069 (0.063)	−0.031 (0.079)	−0.016 (0.070)	−0.064 (0.071)	−0.102 (0.083)	−0.034 (0.062)
Philosophy 2	−0.195 (0.134)	−0.056 (0.170)	−0.120(0.147)	0.019 (0.148)	0.034 (0.174)	−0.058 (0.141)
Philosophy 3	0.087 (0.109)	0.046 (0.137)	0.063 (0.119)	0.119 (0.121)	0.280* (0.139)	−0.037 (0.110)

Note. Standard errors are in parentheses. CK = content knowledge; PCK = pedagogical content knowledge.
*$p < .05$.

pre to post gain for teachers in the face-to-face condition was on average 0.41 points and was well above 1/2 of a *SD* gain. These results suggest teachers in both conditions improved with respect to personal beliefs. Teachers in the online condition had an advantage that was on average 0.11 points, but the difference was not significant at the .05 level. In contrast, the pre to post gain in impersonal beliefs about personal efficacy in teaching environmental science for teachers in the online condition was on average −0.30 points and was larger than 1/2 of a *SD* loss. The pre to post gain for teachers in the face-to-face condition was on average −0.24 points and was larger than 1/3 of a *SD* loss. These results suggest teachers did not improve with respect to impersonal beliefs in either condition. Teachers in the online condition were at a slight disadvantage that was on average −0.06 points, but the difference was not significant at the .05 level. In the regression models, we regressed posttest scores on the treatment conditions and controlled for pretest scores, teacher background (e.g., science coursework completed), teaching certificate in environmental science, and years of teaching experience in environmental science. These results suggest no significant differences between conditions with respect to personal and impersonal efficacy beliefs. Results are reported in Table 2.

Next, we examined differences across conditions in teacher beliefs about teaching environmental science. We analyzed six outcome variables: attitudes toward teaching environmental science, preparedness with respect to CK and PCK in environmental science, as well as beliefs about traditional teaching practices, inquiry teaching practices, and investigative culture. The pre to post gains in all six variables were positive in the online condition. Similarly, the pre to post gains were positive in the face-to-face condition except for investigative culture. The differences in these means between the two conditions were positive except for inquiry practices and attitudes toward teaching environmental science. However, none of the differences in these means were significant. We ran regression models where we regressed the posttest scores on the treatment binary variable as well as pretest scores, teacher background (e.g., science coursework completed), teaching certificate in environmental science, and years of teaching experience in environmental science. We also included covariates about the philosophy of teaching science (e.g., traditional vs. constructivist orientation). Overall, estimates produced from regression models suggested no significant differences between face-to-face and online conditions for all six dependent variables. That is, we were unable to detect any treatment effects with respect to teacher beliefs in both conditions (Table 3).

Another set of teacher outcomes of interest related to teachers enacting the making connections component of the curriculum materials (RQ2). Overall, we examined treatment differences in 23 variables, detecting only five significant differences between conditions. Teachers in the online condition were more likely to read through and discuss essential questions and lesson overviews when introducing a lesson than the face-to-face teachers. Online teachers were less likely to substitute components of a lesson than the face-to-face teachers. In addition, online teachers were more likely to refer to the text when teaching than the face-to-face teachers. We analyzed teacher outcomes by exploring 16 variables with respect to decision making and eight different variables for using technology when implementing the curriculum. We did not detect any significant differences between conditions.

Student Outcomes

The main student outcome of interest (RQ3) was gain in end-of-unit environmental science scores. Students in the online

Table 4. Two-Level Estimates of Student Learning Outcomes.

Variables	Estimate	*SE*
Online versus face-to-face	−0.400	0.757
Pretest	0.427*	0.032
Overall GPA	0.724*	0.275
Science GPA	0.109	0.159
Mathematics GPA	0.233	0.143
Female	0.134	0.254
Black	−0.979	0.552
Hispanic	0.230	0.478
Asian	−0.410	0.946
Other race	−0.339	0.475
Upper versus lower grades	0.435	0.526
Teacher pretest CK scores	0.276	0.180
Course work in biological sciences	−0.101	0.182
Course work in earth sciences	−0.028	0.125
Course work in environmental sciences	0.041	0.105
Certification in environmental sciences	1.417	0.881
Years of teaching experience in environmental science	−0.069	0.054
Teacher impersonal efficacy beliefs	−1.323	1.229
Teacher personal efficacy beliefs	1.945	1.255
Urban school	1.883	1.055
Suburban school	0.726	0.739
Percent of students eligible for free/reduced price lunch	−2.141	1.814
Cohort 1	0.590	1.717
Cohort 2	0.118	0.768

Note. CK = content knowledge.
$*p < .05$.

condition gained from pre to post three points on average on the environmental science scale, a gain larger than 1/2 of a *SD*. In the face-to-face condition, students also gained two points from pre to post on average in the environmental science scale, a gain smaller than 1/2 of a *SD*. That is, students in both conditions improved their science scores. However, the difference in means between the two conditions, although positive, favoring the online condition, was not significant at the .05 level.

The results of the two-level analysis are summarized in Table 4. Once all covariates were included in the model the main independent variable estimate was negative but insignificant at the .05 level, indicating no differences in environmental science scores between students in the two conditions. The only two covariates that were positive and significant were the pretest science scores and the general GPA. All other variables were not different from zero. These results indicate that although students benefited from both treatment conditions, the overall treatment effect was not different from zero and thus the benefit was similar.

Discussion

This is among the first random assignment studies of teacher learning from PD with respect to PD modality, and one of a growing number of studies of teacher learning that includes measures relating teacher learning, classroom practice, and student learning, a logic chain that is important for the field (Borko, 2004; Desimone, 2009; Fishman et al., 2003; Loucks-Horsley & Matsumoto, 1999). This study and others that examine the effects of PD modality using randomized designs (Fisher et al., 2010; Masters et al., 2010; Powell et al., 2010) are important steps toward broadening and deepening the empirical base for understanding how online PD fits into the broader landscape of teacher learning initiatives (Dede et al., 2009).

Something this study shares with earlier studies (Fisher et al., 2010; Masters et al., 2010; Powell et al., 2010) is a finding of no significant differences between conditions. This might be interpreted as no finding at all, but we think it is meaningful and important. Why should there be a difference in learning outcomes for teachers or students just because of PD modality? Although the content of our PD was designed to be the same across both conditions, the affordances differed and thus offered slightly different ways to engage each group of teachers. Affordances of the online PD, such as proximity to practice, the ability to reflect on prior or proximal practice as part of PD because the PD happens during the school year and not during the summer before, and the ability to move at one's own pace, to work in small chunks of time, or to review individual PD lessons may balance out the affordances of the face-to-face PD, which included greater collegiality, sharing of information among teachers, and emergent discussions among participants gathered in the same physical location for an extended period of time.

The issue of total "contact hours" for PD also warrants some consideration. If the average number of hours teachers spent on online PD was less than the total number of hours spent by teachers in the face-to-face condition, might this difference in contact hours impact teacher learning and, subsequently, classroom practice and student outcomes? Our answer is that the tremendous range of time spent in the online PD by teachers represents a valuable affordance of online PD: You only have to spend as much time as you need. There is no "break time" built into online PD; one moves as fast as one wants to. Because we found no relationship between amount of time teachers spent on online PD and our outcome measure, we conclude that teachers who completed online PD more rapidly required less time to benefit from materials. Online PD does not have to be "one size fits all" in terms of participation.

Our study design establishes a linkage between PD and learning outcomes for teachers and students, and in doing so further demonstrates that desirable outcomes need not be contingent on PD modality. In comments in *The Chronicle of Higher Education* on a recent study of student learning in hybrid online/face-to-face university courses that also found no significant differences (Bowen, Chingos, Lack, & Nygren, 2012), one of the study authors mused that the findings might appear to be "a bland result," but in fact were important: "One of the responses most frequently raised in efforts to experiment with [online] teaching is that it will expose students to risk . . . The results of [the] study show that such worries are overblown" (Bowen, as quoted in Mangan, 2012). The risk referred to is that students will have a subpar learning experience online. Arguments about the "risks" of online teacher PD have shared this concern—that efficiency gains of moving online come at the cost of learning or quality.

Given that we (and others) find no appreciable differences in learning outcomes between modalities, one might turn to other factors when considering whether online or face-to-face PD makes the most sense in a given teacher learning context. The idea that online PD is more cost-effective than face-to-face is widely held among policy makers and school leaders (Dede et al., 2009; Kleiman, 2004; Means et al., 2009), though in reality this is highly dependent on context. Although a full cost–benefit analysis is beyond the scope of this paper (for a discussion of the subject, see Odden, Archibald, Fermanich, & Gallagher, 2002), our observations in designing and delivering PD in both modalities is that many costs are similar, especially in the start-up design and development phase. Overall, it is important to keep in mind that principled, carefully designed, and thoughtfully delivered PD will always have significant costs associated with it, no matter what the delivery medium.

If the costs of developing and deploying PD are similar across modalities, major differences in the cost–benefit of online versus face-to-face PD will likely be dependent on other contextual issues, largely related to the nature of the delivery of the PD program. For instance, if one is designing PD for use over a limited period of time by teachers in one geographic area, face-to-face is likely to be more cost-effective than online. However, if one intends to use the same PD materials over and over again, or if the teachers are geographically distributed, then online is likely to be more efficient. In our case, we were designing PD to facilitate teacher adoption of new curriculum materials, which will hopefully be adopted and used by teachers who are geographically distributed, because there is rarely more than one environmental science teacher at any particular school. Therefore, online PD seems to be the cost-effective solution for the adoption of this curriculum, especially when the outcomes can be shown to be equivalent to the "business as usual" model for curriculum adoption, which has long employed face-to-face workshops where teachers must travel to a regional meeting, or the publishers send representatives over long distances to work with individual districts.

Conclusion

We conclude with caution and hope. We do not believe that the PD described in this study, whether online or face-to-face, is representative of *all* PD. Yet there is a tendency, especially among educational decision makers, to treat "online" PD as if it represents a particular approach, rather than a delivery vehicle. There is a need for studies that examine many different designs for PD: "online" and "face-to-face" are no more descriptive in the end than "workshop" in terms of understanding the nature of the teaching and learning activities contained within the PD. This study focused on a common challenge for teachers in science: learning to use new curriculum materials. We encourage readers to be mindful of differences between our PD designs and the ones being considered for other applications. There will be contexts where either online or face-to-face PD is more desirable, for reasons of cost, location, or content. This study represents a step toward a more balanced examination of PD modality to inform policy and practice for teacher learning.

Acknowledgments

We thank Terri Ridenour, Meridith Bruozas, Anna Switzer, Kirsten Mawyer, Vicki James, Colleen Petersen, Yi-Ling Hsaio, Jared Eno, and PreScillia Fleming for support and assistance with recruitment, programming, data processing, and data analysis, and the anonymous reviewers for their feedback and guidance on the manuscript. Thanks also to the many teachers who contributed to this research and to It's About Time publishing for assistance with teacher recruitment.

Authors' Note

The opinions reported in this work are those of the authors and not of the funding agencies or their institutions.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Coauthor Daniel C. Edelson is the lead author of the curriculum program that was the focus of the professional development studied in this paper. As such, he receives royalties on sales of the curriculum. However, the authors assert that there is no conflict of interest with the research because Edelson's financial interest is not aligned with the possible outcomes of the study, and he did not participate in data collection or analysis for the study.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by the National Science Foundation under Grant ESI-0455582, and by the National Geographic Education Foundation.

References

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—Or might be—The role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, *25*(9), 6-8.

Barab, S. A., & Luehmann, A. L. (2003). Building sustainable science curriculum: Acknowledging and accommodating local adaptation. *Science Education*, *87*(4), 454-467.

Becker, H. J. (1999). *Internet use by teachers: Conditions of professional use and teacher-directed student use* (Report No. 1). Irvine: Center for Research on Information Technology and Organizations, University of California, Irvine. Retrieved from http://www.crito.uci.edu/TLC/findings/Internet-Use/startpage.htm

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3-15.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2012). *Interactive learning online at public universities: Evidence from randomized trials*. New York, NY: Ithaka S+R.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2012). Great teaching: Measuring its effects on students' future earnings. *EducationNext*, *12*(3). Retrieved from http://educationnext.org/files/ednext_20123_Friedman.pdf

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

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2), 119-142.

Crawford, B. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching*, *37*, 916-937.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 2-14.

Dede, C., Ketelhut, D. J., Whitehouse, P., & McCloskey, E. (2009). A research agenda for online teacher professional development. *Journal of Teacher Education*, *60*(1), 8-19.

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199.

Desimone, L., Porter, A., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81-112.

Edelson, D. (2001). Learning-for-use: A framework for the design of technology-supported inquiry activities. *Journal of Research in Science Teaching*, *38*(3), 355-385.

Edelson, D., Schwille, K., Bruozas, M., Lach, M., Taber, M., Gordin, D., & Tarnoff, A. (2005). *Investigations in environmental science: A case based approach to the study of environmental systems*. Armonk, NY: It's About Time.

Fisher, J. B., Schumaker, J. B., Culbertson, J., & Deshler, D. D. (2010). Effects of a computerized professional development program on teacher and student outcomes. *Journal of Teacher Education*, *61*(4), 301-312.

Fishman, B., Marx, R., Best, S., & Tal, R. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, *19*(6), 643-658.

Frechtling, J. A., Sharp, L., Carey, N., & Vaden-Kiernan, N. (1995). *Teacher enhancement programs: A perspective on the last four decades*. Retrieved from http://www.ehr.nsf.gov/ehr/rec/pubs/eval/tep/tep.htm

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., … Sztejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement* (NCEE No. 2008–4030). Washington, D.C.: Institute of Education Sciences.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945.

Ginsberg, A., Gray, T., & Levin, D. (2004). *Online professional development for mathematics teachers: A strategic analysis*. Washington, DC: National Center for Technology Innovation, American Institutes for Research.

Kao, C., & Tsai, C. (2009). Teachers' attitudes toward web-based professional development, with relation to Internet self-efficacy and beliefs about web-based learning. *Computers & Education*, *53*(1), 66-73.

Kleiman, G. M. (2004). *Meeting the need for high quality teachers: E-Learning solutions*. Retrieved from http://www2.ed.gov/about/offices/list/os/technology/plan/2004/site/documents/Kleiman-MeetingtheNeed.pdf

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, *42*(2), 7-19.

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*(1), 1-32.

Kubitskey, B., Fishman, B., & Marx, R. W. (2002). Professional development, teacher learning, and student learning: Is there a connection? In P. Bell, R. Stevens, & Y. Satwicz (Eds.), *Proceedings of the International Conference of the Learning Sciences* (pp. 229-236). Mahwah, NJ: Lawrence Erlbaum.

Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks, CA: Corwin Press.

Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics*, *99*(5), 258-271.

Mangan, K. (2012, May 22). Study shows promise and challenges of "hybrid" courses. *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/blogs/wiredcampus/study-shows-promise-and-challenges-of-hybrid-courses/36350

Masters, J., Magidin deKramer, R., O'Dwyer, L. M., Dash, S., & Russell, M. (2010). The effects of online professional development on fourth grade English language arts teachers' knowledge and instructional practices. *Journal of Educational Computing Research*, *43*(3), 355-375.

McGee, S., & Brazdil, L. C. (2011, April). *The influence of teacher grading practices on course performance and high stakes test performance*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: U.S. Department of Education.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*(3), 237-257.

Odden, A., Archibald, S., Fermanich, M., & Gallagher, H. A. (2002). How to figure the cost of professional development. *Journal of Staff Development* *23*(2), 53-58.

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

O'Dwyer, L. M., Masters, J., Dash, S., Magidin deKramer, R., Humez, A., & Russell, M. (2010). *e-Learning for educators: Effects of on-line professional development on teachers* *and their students: Findings from four randomized trials*. Retrieved from http://www.bc.edu/research/intasc/PDF/EFE_Findings2010_Report.pdf

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology*, *102*(2), 299-312. doi:10.1037/a0017763

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis*. Newbury Park, CA: SAGE.

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*(6), 625-637.

Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, *39*, 369-393.

Santagata, R., Kersting, N., Givvin, K. B., & Stigler, J. W. (2011). Problem implementation as a lever for change: An experimental study of the effects of a professional development program on students' mathematics learning. *Journal of Research on Educational Effectiveness*, *4*(1), 1-24. doi:10.1080/19345747.2010.498562

Scher, L., & O'Reilly, F. (2009). Professional development for K-12 math and science teachers: What do we really know? *Journal of Research on Educational Effectiveness*, *2*(3), 209-249. doi:10.1080/19345740802641527

Schlager, M. S., Fusco, J., & Schank, P. (2002). Evolution of an online education community of practice. In K. A. Renninger & W. Shumar (Eds.), *Building virtual communities: Learning and change in cyberspace* (pp. 129-158). New York, NY: Cambridge University Press.

Smith, P. S., Banilower, E. R., McMahon, K. C., & Weiss, I. R. (2002). *The national survey of science and mathematics education: Trends from 1977 to 2000*. Chapel Hill, NC: Horizon Research. Retrieved from http://2000survey.horizon-research.com/reports/trends.php

Squire, K. D., MaKinster, J. G., Barnett, M., Leuhmann, A., & Barab, S. A. (2003). Designed curriculum and local culture: Acknowledging the primacy of classroom culture. *Science Education*, *87*(4), 468-489.

Tal, T., Krajcik, J. S., & Blumenfeld, P. C. (2006). Urban schools' teachers enacting project-based science. *Journal of Research in Science Teaching*, *43*(7), 722-745.

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

Author Biographies

Barry Fishman, PhD, is associate professor in the School of Education and School of Information at the University of Michigan. His research focuses on teacher learning with technology and the development of scalable innovation through design-based implementation research.

Spyros Konstantopoulos, PhD, is associate professor of measurement and quantitative methods at the department of counseling, educational psychology, and special education at the College of Education at Michigan State University. His research interests include teacher and school effects, class size effects, program evaluation, research design, and meta-analysis.

Beth W. Kubitskey, PhD, is associate professor of physics and astronomy and assistant director of the W. K. Kellogg Foundation Woodrow Wilson Michigan Teaching Fellowship at Eastern Michigan University. Her research interests include designing, implementing, and studying inservice and preservice STEM teacher education and its impact on student learning.

Richard Vath, MA, is a doctoral candidate in the Combined Program in Education and Psychology at the University of Michigan. His research interests include learning analytics in online learning environments as well as the impact of professional development opportunities on physician educators' scholarly thinking and practice.

Gina Park, MS, is a doctoral candidate in the Combined Program in Education and Psychology at the University of Michigan. Her research interest is on the impact of teacher learning opportunities on changes in teachers' knowledge, beliefs, and teaching practice.

Heather Johnson, PhD, is an assistant professor of the practice of science education in the Department of Teaching and Learning at Vanderbilt University. Her research focuses on developing supports for teacher learning and exploring how these supports affect teacher practice and ultimately student learning.

Daniel C. Edelson, PhD, is vice president for education at the National Geographic Society. In this role, he oversees the Society's outreach to educators and its efforts to improve geographic and geoscience education in the United States and abroad.